AALTO UNIVERSITY SCHOOL OF SCIENCE

Department of Industrial Engineering and Management

Institute of Strategy

Välimaa, Norma

STRATEGIC DEMAND MANAGEMENT IN DIAGNOSTIC ELECTRONEUROMYOGRAPHY TESTING: ANALYSIS OF FACTORS INFLUENCING EFFECTIVENESS OF TESTS

Master's Thesis

Espoo 27.5.2015

Supervisor:	Professor Jens Schmidt
Instructors:	D.Sc. (Tech.), Assistant professor Antti Peltokorpi
	M.D., PhD, Adjunct professor Erika Kirveskari

AALTO UNIVERSITY SCHOOL OF SCIENCE
Degree Programme in Industrial Engineering and Management

ABSTRACT OF THE MASTER´S THESIS

<table>
<tr><td colspan="3">Author:
Norma Välimaa</td></tr>
<tr><td colspan="3">Subject of the thesis:
Strategic demand management in diagnostic electroneuromyography testing: Analysis of factors influencing effectiveness of tests</td></tr>
<tr><td>Number of pages:
109</td><td>Date:
27.5.2015</td><td>Library location:
TU</td></tr>
<tr><td colspan="2">Professorship:
Strategic management</td><td>Code of professorship:
TU-91</td></tr>
<tr><td colspan="3">Supervisor:
Antti Peltokorpi, D.Sc. (Tech.), Assistant professor
Erika Kirveskari M.D., Ph.D., Adjunct professor</td></tr>
<tr><td colspan="3">Instructors:
Jens Schmidt, Professor</td></tr>
</table>

The purpose of this study is to explore effectiveness of diagnostic tests and demand management strategies to increase this effectiveness. The study focuses on significance of electroneuromyography (ENMG) tests and explores the factors affecting the significance of ENMG tests, which affect the strategy to be selected. The assumption in this study is that a diagnostic test is insignificant if that test does not affect patient diagnosing, treatment plan decision-making, or provide any additional value to referring physician. The case company is a Finnish public utility, HUS Medical Imaging Center, providing diagnostic tests to the largest hospital districts in Finland; Hospital District of Helsinki and Uusimaa (HUS).

The literature review focuses on effectiveness of healthcare services. Within healthcare, effectiveness is traditionally defined through patient outcome, and we further review prior studies' definitions of effectiveness among diagnostic testing, specifically. Since the purpose of the focal organization is to reduce the number of insignificant tests with demand management strategies, the literature review also focuses on defining insignificant tests and explores demand management strategies.

The empirical part of this study uses case study approach and survey research methodologies. We explore local opinions and factors contributing the studied phenomenon with the case study method, and employ the survey research method for data gathering. Surveys were conducted to two groups of physicians: neurophysiologists and referring physicians. In addition, three interviews were conducted to experts in the field to gather support for the findings of this study. Furthermore, performance data concerning clinical neurophysiology laboratories were gathered from years 2011-2014. Survey results and performance data were statistically tested with chi-square tests, ordinal logistic regression model and Student's t-test.

The results suggest that roughly 11% of the ENMG tests conducted in HUS clinical neurophysiology (CN) laboratory in Meilahti are insignificant. Interestingly, referring units differ in significance, as primary care and physiatry outpatient clinics request less significant ENMG tests than other secondary care outpatient clinics. Furthermore, the findings of this study support the earlier results that referral queries contribute to the value of diagnostic tests. Therefore, we suggest that units providing diagnostic test services highlight the importance of clear referral queries, educate referring physicians, and use different demand management strategies for different referring quarters. Finally, further actions are suggested for improving the effectiveness of ENMG testing and employing demand management strategies.

Keywords: electroneuromyography, effectiveness, demand management strategy, diagnostic testing, healthcare, significance	Publishing language: English

AALTO UNIVERSITY SCHOOL OF SCIENCE
Degree Programme in Industrial Engineering and Management

ABSTRACT OF THE MASTER´S THESIS

Tekijä:
Norma Välimaa

Työn nimi:
Elektroneuromyografia tutkimusten strateginen kysynnänhallinta: analyysi diagnostisten testien vaikuttavuuteen vaikuttavista tekijöistä.

Sivumäärä:	Päiväys:	Työn sijainti:
109	27.5.2015	TU

Professuuri:	Koodi:
Strateginen johtaminen	TU-91

Työn valvojat:
Antti Peltokorpi, TkT, Professori (professuuri: Rakentamisen tuotantotalous)
Erika Kirveskari, LT, Apulaisprofessori

Työn ohjaaja:
Jens Schmidt, Professori

Tutkimuksessa on tarkoituksena selvittää diagnostisten testien vaikuttavuutta, ja millä kysynnän hallinnan strategioilla vaikuttavuutta voitaisiin parantaa. Tämä tutkimus keskittyy selvittämään elektroneuromyografiatutkimusten (ENMG) merkitsevyyttä, mikä lopulta vaikuttaa kysynnän hallinnan strategian valintaan. Oletuksena työssä on, että diagnostinen tutkimus on merkityksetön, jos se ei vaikuta potilaan diagnosointiin, hoitopäätöksen tekoon, tai ei tuo lisäarvoa lähettävälle lääkärille. Kohdeyritys on suomalainen HUS Kuvantamisen liikelaitos, joka tuottaa kuvantamisen palveluita Suomen suurimmalle sairaanhoitopiirille: Helsingin ja Uudenmaan sairaanhoitopiirille.

Kirjallisuuskatsaus keskittyy terveydenhuollon vaikuttavuutta tutkivaan kirjallisuuteen. Terveydenhuollon vaikuttavuutta tarkastellaan usein potilaan hoitotuloksen kautta, ja tämä tutkimus kartoittaa, miten vaikuttavuutta mitataan diagnostisten testien osalta. Muutama aiempi tutkimus on esittänyt mittareita tähän tarkoitukseen. Koska tutkimuksen tarkoituksena on vähentää merkityksettömien tutkimusten määrää kysynnän hallinnan strategioiden avulla, tarkastellaan kirjallisuuskatsauksessa tarkemmin sekä merkityksettömän diagnostisen tutkimuksen määritelmää ja erilaisia kysynnän hallinnan strategioita.

Metodeina käytämme case- ja kyselytutkimusmenetelmiä. Case-tutkimusmenetelmillä tutkitaan paikallisia mielipiteitä ja tekijöitä, jotka vaikuttavat tutkittuun ilmiöön. Mittaustiedon kerääminen suoritettiin kyselytutkimuksilla. Kyselyt suoritettiin sekä lähettäville lääkäreille että neurofysiologeille. Lisäksi suoritettiin kolme haastattelua alan asiantuntijoille löydösten tueksi. HUS:n kliinisen neurofysiologian (KNF) yksiköistä kerättiin suoritetietoja vuosilta 2011–2014. Data analysoitiin tilastollisin menetelmin khiin neliötesteillä, ordinaalisella regressiomallilla ja Studentin t-testillä.

Tuloksemme osoittavat, että noin 11 % HUS:n KNF-laboratoriossa Meilahdessa suoritetuista ENMG-tutkimuksista on merkityksettömiä. Tutkimuksen mukaan lähettävä taho vaikuttaa tutkimuksen merkitsevyyteen: perusterveydenhuolto ja fysiatrian poliklinikka tilaavat merkityksettömämpiä ENMG-tutkimuksia kuin muut erikoissairaanhoidon poliklinikat. Lisäksi tämän tutkimuksen tulokset vahvistavat aiempien tutkimusten tuloksia koskien lähetteen kysymyksenasettelun tärkeyttä merkitsevyyden kannalta. Tämän vuoksi suosittelemme, että diagnostisia testejä tuottavat yksiköt korostaisivat selkeän lähetteen merkitystä, järjestäisivät koulutusta lähettäville tahoille, sekä käyttäisivät eri kysynnänhallintakeinoja eri lähettäville tahoille. Lopuksi ehdotamme muita toimia, joilla voidaan nostaa ENMG-testien merkitsevyyttä sekä ottaa käyttöön kysynnän hallinnan strategioita.

Avainsanat: elektroneuromyografia, vaikuttavuus, kysynnän hallinnan strategia, diagnostinen testaus, terveydenhuolto, merkitsevyys	Julkaisukieli: Englanti

Acknowledgements

The process of writing this master's thesis has been challenging, but still rewarding and above all educational. The topic of effectiveness within healthcare is timely and very interesting, giving many new ideas and ways to improve the current healthcare system.

At this point, I want to sincerely thank my instructor *Erika Kirveskari* from HUS Medical Imaging Center. I am extremely thankful for this most suitable and interesting topic she gave me. Furthermore, her support throughout the process, long experience within the field of clinical neurophysiology and thus her valuable insight into the research was what made this master's thesis possible. Our discussions have offered the most important inspiration into the work. Inspiring has also been her strong positive and supportive attitude towards this thesis. Working with Erika has positively influenced me, and I really appreciate the time and effort she put into instructing my work.

In addition, great thanks to my instructor *Antti Peltokorpi* from Aalto University, who provided insight into the analyses of this study, as well as helped to keep strong industrial engineering and management perspective throughout the thesis. The structure and overall input of this study was improved significantly according to his comments. Furthermore, I want to thank professor *Jens Schmidt* for evaluating and supervising this master's thesis and giving valuable insight from the strategic perspective.

Special thanks to *HUS Medical Imaging Center* for letting me to study and learn more of this topic within their diagnostic processes and giving me the information needed for this master's thesis. I also thank all physicians who took the time to provide information for this thesis.

Finally, I want to genuinely thank my dear family, who has supported me in many ways throughout my life, the 18 years in schooling and to pursue this master's thesis. Furthermore, my dear boyfriend and friends also deserve very special thanks for valuable comments and supporting and encouraging me.

At the moment, being very close to the Master's Degree at Aalto University, I remember the six years spent here with great gratitude. These six years have exceeded my expectations, and I am confident that I will move forward with much more valuable knowledge and capabilities that I ever hoped for. Furthermore, I revere the times spent with the brilliant people from the Department of Industrial Engineering and Management from Aalto University. Especially, the time spent in Prodeko events reminds me of the fellow students that have made these years exceptional and memorable.

Helsinki May 27, 2015

Norma Välimaa

Table of Contents

List of figures

List of tables

1. Introduction

1.1 Background

The effectiveness of healthcare services has become an ever more important issue during the past years, as the population structure is shifting in most western countries; the working ages are diminishing while at the same time, the population is ageing and demand for healthcare services is increasing (Rechel et al. 2013; Timonen & Kautto 2014). This leads to cost pressures and lack of resources. Furthermore, physicians rely more and more on diagnostic tests rather than on pure clinical examination of a patient due to uncertainty (Fuller 2005), causing an increase in the demand for diagnostic tests. Further, diagnostic tests have become more accessible, faster and inexpensive which makes a larger number of requests possible compared to the past (Fuller 2005). Besides the increased number of patients, reasons for the increased demand identified in the literature include physicians' inexperience, uncertainty, lack of awareness of the current guidelines, pressure from patients who nowadays have more information about their condition and therefore may demand certain tests, medico-legal reasons, and the current culture among the new generation of physicians of relying mainly on instrumental evidence (Di Fabio et al. 2013; Sood et al. 2007; Solomon et al. 1998). All in all, diagnostic imaging, and especially advanced imaging and diagnostic tests, contribute to healthcare costs, thus their more appropriate use will improve the whole healthcare system (Blackmore et al. 2011). Because the current situation of increasing demand and limited resources, as well as the lack of effectiveness research in the healthcare sector, it is interesting and valuable to research this phenomenon and gather information from literature concerning effective healthcare services.

The effectiveness of a diagnostic tool is hard to measure. This is because diagnostic tests do not impact the patient outcome directly, in contrast to i.e. surgical treatment, but rather enable the formulation of appropriate diagnoses and treatment plans (Neumann & Weinstein, 2010.) Thus, the diagnostic tool functions as a small but crucial part in the patient's treatment pathway. According to the literature review conducted as part of this study, this field of research seems to be underdeveloped. Effectiveness studies in healthcare commonly use a method where the outcomes of two medical tools or treatment methods are compared (Gazelle et al. 2011; Lee et al. 2010), or comparing costs – called cost-effectiveness analysis (CEA) (Singer, M.E. & Applegate, K.E., 1996). Contrastingly, this study explores and measures effectiveness through the volume of insignificant examinations.

The aim of this paper is to study the effectiveness of electroneuromyography (ENMG) tests. An ENMG test is an electrodiagnostic test which examines electrical functioning of nerves and muscles. According to my knowledge, the Finnish public healthcare sector has not conducted such effectiveness studies before. However, effectiveness of services is currently listed as one of the key strategies in the hospital district of Helsinki and Uusimaa (HUS) (Helsingin ja Uudenmaan Sairaanhoitopiiri 2012), which is the largest hospital district in Finland. Internationally, few studies have examined inappropriate ENMG test requests within the field of clinical neurophysiology (Cocito et al. 2006; Di Fabio et al. 2013; Mondelli et al. 2014; Podnar 2005; Rigler & Podnar 2007; Mondelli et al. 1998). These studies have been conducted in Turkey, Slovakia, the U.S., and Italy (Di Fabio et al. 2013; Mondelli et al. 1998; Mondelli et al. 2014; Cocito et al. 2006; Podnar 2005; Shepherd 2010; Karadag et al. 2014), countries that have very dissimilar healthcare systems from Finland. Thus, this study is the first of its kind in a welfare country such as Finland.

The effectiveness of diagnostic test relates closely to the definition of significant tests and appropriate test requests (Fryer & Smellie, 2013). The research about the criteria and characteristics of inappropriate diagnostic tests is wide (van Walraven & Naylor 1998). The identification of inappropriate tests is important, because it provides guidance to requesting; which test requests should be reduced in order to make healthcare services more effective. In the present study, the concept of significance means that the test is judged by a physician to be significant if it helps the physician to either diagnose the patient or decide the treatment for the patient. In this study, the concept of significance does not concern technical aspects, such as whether the test functions technically as it should or if the test is performed technically correctly. The definition of an inappropriate test request and an insignificant test are described in more detail in the sub-chapter 1.5.

The definition of an inappropriate test brings us to the strategic goal of a clinical neurophysiology (CN) laboratory: they need to manage demand by prioritizing test requests that are relevant and appropriate. Demand management studies have been widely conducted in the context of laboratories in general (Fryer & Smellie 2013; Janssens 2010), but not in the field of CN laboratories specifically. Demand management and demand control are separate concepts: demand control methods aim to reduce the number of requests, while demand management mainly strives to ensure appropriate requests. Thus, demand management has an inherent quality aspect at its core, and does not necessarily lead to fewer, but more appropriate, requests. (Fryer & Smellie 2013, p.62.) Overall, with a valid demand management strategy, a CN laboratory is able to ensure sufficient resources to meet the critical demand while keeping a high level of quality of care.

The present study is initiated by the public utility *HUS Medical Imaging Center's* CN laboratory in Meilahti, operating in the Hospital District of Helsinki and Uusimaa (HUS), Finland. The present research observes HUS as a whole, which includes eight CN laboratories. The demand of ENMG tests in HUS has increased significantly during the last few years, resulting in increasing costs. An ENMG test is labor intensive, requiring 30 to 60 minutes of specialist time per test. During the last few years, HUS has tried to reduce costs through reducing the number of outsourced physicians, and shorten waiting times through Lean methodologies (Womack et al. 1992) and extra tests conducted outside office hours. However, these efforts have not been sufficient to tackle the problem.

Although much has been made in order to increase the ENMG test supply, it is not enough, as demand is growing constantly. To this point also the effectiveness of healthcare services plays a huge role in providing the sufficient supply. In a situation where we have to balance between limited resources and great number of patients in need for the services, we minimize the waste of resources and ensure the quality of care by effective services. By studying the effectiveness of a healthcare service we might be able to prioritize the demand and focus the service to those patients who present actual and real need for the service. It is essential to provide the maximal health benefit with the limited resources we have. The effectiveness of healthcare services and the usage of appropriate demand management strategies make both cost savings and more efficient resource allocation possible, which benefit both the service orderer and producer.

1.2 Objective of the study

The aim of this research is to gain understanding of how effective electroneuromyography (ENMG) tests are in the light of diagnostics and treatment plan decision-making, and then to develop strategies to increase the effectiveness and the value of these diagnostic tests for end-customer. These strategies are mostly meant for managing the demand by controlling and guiding the behavior of referring physicians. First, this study identifies and then tries to reduce the number of inappropriate ENMG test requests and insignificant ENMG tests by proposing various demand management strategies. These strategies are explored in more detail in the literature review section. Additionally, the concept and characteristics of an insignificant ENMG test is important to identify and determine in order to select appropriate strategies. The concept of insignificant diagnostic test is firstly clarified in the literature review and then based on this search, similarities and differences are searched from the gathered data by analyzing it with qualitative and quantitative methods.

The evaluation of effectiveness of ENMG tests is subjective, thus this study will evaluate it from three different viewpoints: 1) referring physician, 2) neurophysiologist and 3) experienced referring clinical specialist. In contrast to prior studies this study explores the significance of ENMG test from both sides: test orderer and test producer.

The postulate is that an insignificant ENMG test does not affect patient treatment plan or diagnosis and that the significance is highly dependent on the formulation and focus of referral query. Before performing this research a few hypotheses concerning the characteristics of insignificant ENMG test were concluded. The first hypothesis is that primary care order ENMG tests with unfocused question setting, while more specialized outpatient clinics within secondary care, mainly formulate more focused referral queries. The second hypothesis is that unfocused referral questions result in insignificant ENMG tests. This is due to the assumption that general and unfocused referral queries force neurophysiologist to perform more extensive ENMG tests, because neurophysiologist has to examine several limbs in order to find the part of body that causes the symptoms.

In order to achieve the goal of the study, the following research questions are set. Answering our main question first requires that we answer sub-questions. The questions are listed below, and the hierarchy between the questions is presented in Figure 1.

- What strategies could be used to reduce the number of insignificant diagnostic tests in healthcare?
  - What are the characteristics of insignificant ENMG test?
    - What are the main factors contributing to ENMG tests' insignificance?
    - Does the opinion about ENMG tests' significance vary between referring physicians and neurophysiologists?
  - What number of ENMG test requests and ENMG tests are insignificant in Meilahti CN laboratory and do referring units differ in this?

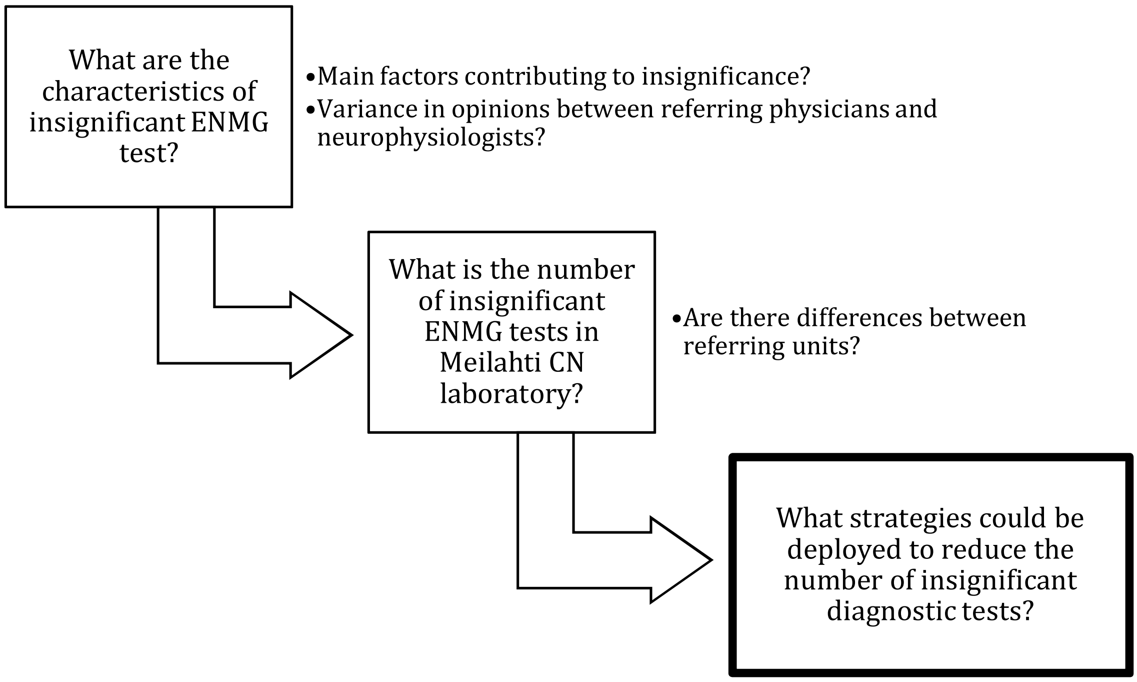

Figure 1 Hierarchy of the research questions

1.3 Research design

This study is a case study focusing on HUS Meilahti CN laboratory and ENMG tests utilizing survey research approach. First, a academic literature review is compiled in order to explore demand management strategies, previous effectiveness studies and characteristics of an insignificant diagnostic test. Second, the data is collected and analyzed. The main data is collected via two surveys made for neurophysiologists and referring physicians. Also, three interviews are conducted to clinical specialists experts in diagnostic test requesting to support the findings of survey data analysis. Performance data of CN laboratories within HUS area is gathered in order to analyze the reasons for increased demand and the differences between the CN laboratories. The data analysis is done with ordinal logistic regression model, chi-square test and Student's t-test. Lastly, the answers to research questions are presented with conclusions by combining the findings from literature to the findings of data analysis. The findings were confirmed by the focal organization of Meilahti CN laboratory.

The literature review is conducted by thorough search of academic articles and other literature. These were search from three main databases among some others; Google scholar, EBSCO e-books (EBSCO-host), Proquest and PubMed. The most relevant keywords used for literature search were among others “EMG testing”, “Effectiveness of healthcare”, “NCS testing”, “inappropriate EMG”, “inappropriate diagnostic testing” and “Demand management strategy”. The literature search was performed in two phases. First, the most relevant articles and other literature found with key words was gathered, and then additional relevant papers were identified in the reference lists of these articles.

1.4 Scope of the study

The focus of the study was decided by public utility of HUS *Medical Imaging Center*. Because all eight CN laboratories within HUS area differ from each other compared by main referring quarters, length of waiting time, and test volumes, it would have been good to gather data from all of CN laboratories and to analyze HUS area as whole. However, as a master’s thesis, the scope had to be limited. The CN laboratory operating in Meilahti and Laakso hospital was selected as a focus, due to the fact that this CN laboratory is the largest CN unit in HUS area as it has the highest test volume and also the longest waiting times.

This study focuses on the most typical ENMG tests, with the extent of the test ranging from less extensive to more extensive – ENMG-1, ENMG-2 and ENMG-3 tests. Despite of the fact that CN laboratory’s test volumes consist of also numerous other electrodiagnostic tests, the focus of analysis is set to ENMG, because it accounts for 45 % of all electrodiagnostic tests performed and there is a long waiting time to ENMG tests. Furthermore, the number of ENMG tests has increased during the past years and ENMG requires a lot of time from neurophysiologist, thus it binds a large percentage of neurophysiologist resources.

In addition to above scopes, this research analyses Meilahti CN laboratory performance data from year 2011 to year 2014. This outline of years is done in order to keep in the limits of master’s thesis. Moreover, Meilahti CN laboratory put a digital test report system into operation during year 2010, thus the test information is only reliable from year 2011.

This study focused to limited number of referring quarters, with whom a survey was conducted. However, referring units requesting the most ENMG tests annually were selected for this research. A digital version of the survey was sent to all referring quarters attached to ENMG test result statements during November 2014, and a paper version of the same survey was sent to only to 10 most active referral quarters, which request ~50% of Meilahti CN laboratory’s annual ENMG tests performed.

The object of analysis is neurophysiologists' and referring physicians' evaluation concerning the significance of ENMG tests. Exploring the difference of their opinions is one of the focuses of this study. Furthermore, this research provides interesting information about the phenomenon of how the variation between individual referring physicians affects the whole picture of effectiveness and ordering of unnecessary diagnostic tests (Mindemark et al. 2010).

One possibility was to analyze treatment pathways of single patients who have been examined by ENMG test, and then evaluate how ENMG test affected the overall patient outcome. However, treatment pathway analysis is omitted from this research due to time and resource limitations of this study and such evaluation would require a lot of work including selecting control groups etc. in order to determine when ENMG test is significant from the patient outcome point of view.

As this study is highly related to effectiveness research, a more detailed study about the effects of financial incentives would have been interesting to conduct and determine the extent to which financial asymmetry explains the ineffectiveness of our healthcare system. In a welfare country, the Finnish healthcare system does not charge the healthcare service requester (referring physician or patient), but the healthcare cost is covered by municipality. In other words, referring physicians have personal financial motives while requesting diagnostic tests. This creates a financial asymmetry between the service producer and orderer: the service requester has no incentives to reduce the costs and the number of services ordered. Due to the limitations of this study, it was not possible to analyze financial incentives. Nonetheless, the focus of the present study is set to be the patient cases instead of the cost of the test, and whether the ENMG test is helpful for a physician making a diagnosis or treatment plan decision. In other words, this research focuses more on the benefit and value creation of a diagnostic test from the patient case point of view, without taking the cost aspect into account.

1.5 Definitions

The key construct in this research is the *inappropriate ENMG test request* and the *insignificant ENMG test*. In order to understand these concepts, one has to understand what *electroneuromyography (ENMG) test* examines and what resources it requires. The notion of *effectiveness of healthcare services* is important since the aim of this research is to improve effectiveness and measure it through the metrics of *effectiveness*. Finally, while this research will provide understanding of the amount and characteristics of insignificant ENMG tests, the focal organization desires to reduce the number of such tests, thus the notion of *demand management strategy* is important.

Electroneuromyography (ENMG) test – ENMG test is an electrodiagnostic test, which examines the electrical functions of nerves and muscles. ENMG tests are done in clinical neurophysiology (CN) laboratories by neurophysiologists. The duration of one ENMG test varies from 30 to 60 minutes, requiring this time from neurophysiologists to perform, thus the ENMG tests are labor-intensive and expensive. From the patient's perspective, an ENMG test is uncomfortable since it includes needles inserted into muscles and electric stimulation. The most common diagnoses examined by ENMG test are upper limb neuropathies, e.g. carpal tunnel syndrome and lower back radiculopathy. Clinical neurophysiology units within HUS area altogether perform approx. 8 550 ENMG tests annually. The extent of an ENMG test varies between levels 1–3 depending on how many muscles, limbs or clinical questions are examined. Level 1 presents an ENMG test where either 1-4 muscles or one nerve segment is examined. Level 2 presents an ENMG test where one limb or one problem is examined, while level 3 includes examination of many limbs or many clinical questions.

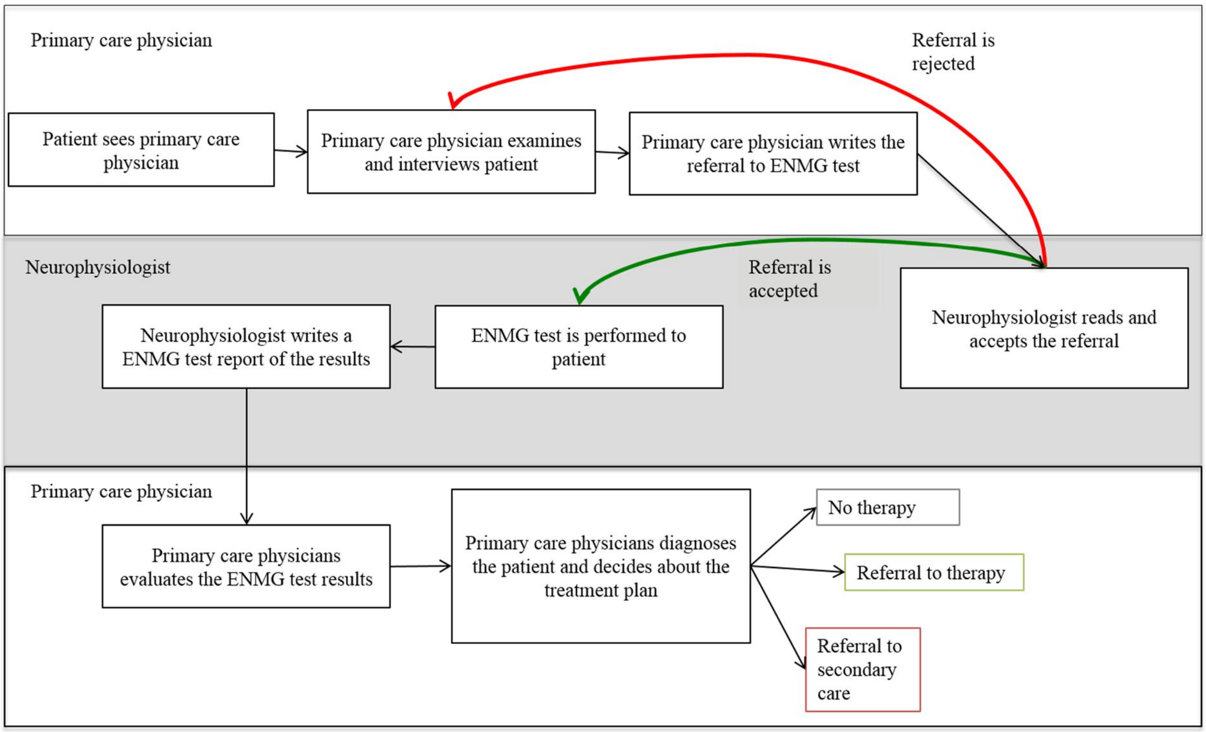

Figure 2 The ordering process of an ENMG test within primary care

Figure 2 presents the ordering process of an ENMG test within primary care. The ordering of an ENMG test differs slightly when ordered by outpatient clinic within secondary care. Firstly, ENMG test referrals from secondary care are often accepted without neurophysiologist's evaluation of test referral, while ENMG test referrals from primary care require neurophysiologist's acceptation before patient is called to ENMG test. Secondly, an ENMG test is ordered for different reasons depending on the referring quarter. Within secondary care, ENMG tests are usually ordered to determine the correct treatment plan, which is either an operation, therapy, additional examinations, observing the healing progress of a patient (control test) or to diagnose a patient. Usually, primary care physicians order ENMG tests to determine a diagnosis and often ENMG tests are ordered for referrals to secondary care.

Effectiveness of healthcare service – An effective healthcare service is a service, which provides maximum health benefit with minimum resources and costs. The common effectiveness metrics are the cost per Quality Adjusted Life Year (QALY), morbidity, and functional status of a patient (Räsänen et al. 2006; Singer & Applegate 2001). However, the effectiveness, health benefit, and value are hard to determine in diagnostic testing due to the indirect nature of its effect on patient outcome (Lee et al. 2010; Neumann et al. 2010). Thus, we measure effectiveness through the *diagnostic thinking efficacy* and *clinical efficacy* measures presented by Fryback & Thornbury (1991). *Diagnostic thinking efficacy* measures effectiveness through the impact of a diagnostic test on diagnosing. *Clinical efficac*y measures the impact on treatment plan decision-making (Fryback & Thornbury 1991.) Diagnostic tests that do not affect either diagnosing or treatment plan decision-making are considered insignificant. As a result, the proportion of insignificant tests in the present study will determine the effectiveness of CN laboratory.

Insignificant ENMG test and inappropriate ENMG test request – Even though the goal of this research is to identify the characteristics of an insignificant ENMG test, there are a few general opinions about which test is considered an insignificant diagnostic tests. Insignificant ENMG tests that are performed to a patient whose diagnosis is obvious already beforehand, without performing ENMG test. Such a diagnosis is called a *self-evident diagnosis*. In addition, insignificant ENMG tests include tests that are ordered without proper clinical examination, or a clinical question or symptoms that cannot be examined by an ENMG test. An insignificant ENMG test does not affect referring physician's decision-making about patient's diagnosis or treatment plan.

An *inappropriate ENMG test request* refers to referrals without clear, focused or relevant referral question setting. An inappropriate ENMG test request may result in insignificant ENMG test if neurophysiologist has to do a wide search, called *screening*, to a patient with the ENMG test device, because the referral does not include a specific symptom or clinical question to be examined. Furthermore, such unfocused referral queries require a lot of neurophysiologists time to interview and clinically examine the patient before performing the test, which additional time is taken from the ENMG examination. Normal results (where the test shows normal function of the muscles and nerves) from ENMG tests are sometimes considered necessary, if the result gives valuable information to the referring physician and helps them in the diagnostics and to plan a treatment for the patient i.e. in pre or post operative situations and when excluding diagnoses. Alternatively, if the ENMG test result is assumed to be normal before performing ENMG test, referred to *self-evident normal* result, then the test is considered to be insignificant. These listed characteristics of insignificant diagnostic test are compared to the final findings of this study.

Demand management strategy – Demand management strategy refers to a selection of approaches used to manage demand and ensure appropriate requesting increasing the effectiveness of the service. It differs from demand control, which refers to explicitly reducing the volume of requesting (Fryer & Smellie 2013.) In the present study, demand management strategy aims at producing appropriate ENMG test that provide the maximum health benefit for the minimum resources spent in producing the ENMG test. Common approaches within diagnostic testing include informing referring physicians about test appropriateness and correct test frequencies (Fryer & Smellie 2013), giving feedback concerning the costs and test usage and emphasizing the importance of dialogue between the service orderer and producer (Gama et al. 1992). Demand management has been reported not to worsen clinical outcomes, even though it in many cases succeeded in reducing the number of test requests (Warren et al. 2010).

1.6 Structure of the master's thesis

This thesis is organized into nine sections (Figure 3). Introduction presents the background and aims of this study. Literature review is divided into three sections, which discuss the topics of demand management strategy, effectiveness of healthcare systems and insignificant diagnostic tests. The research design presents the qualitative and quantitative analysis methods and surveys used in this study. After presenting the research design, the results from surveys, interviews and data gathered from ENMG test report systems is analyzed and presented. The discussion chapter concludes the findings from qualitative and quantitative data analysis and discusses the possible suggestions to the research problem. Validity and reliability of this study are discussed. Finally, the last chapter summaries the findings of this study answering to the research questions presented. In addition, the areas for future research are also discussed in the last chapter.

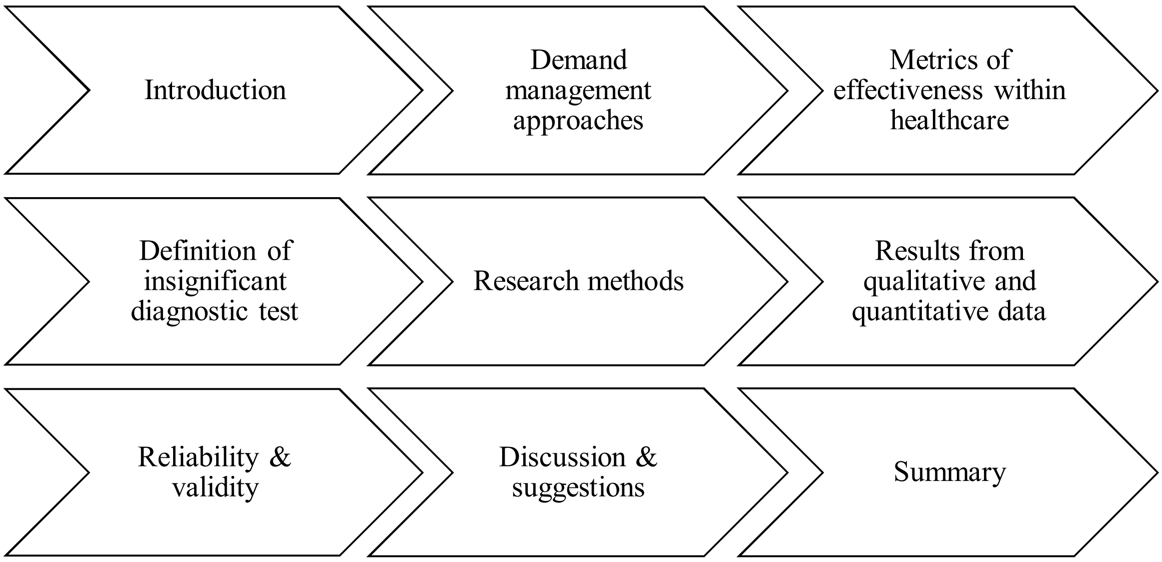

Figure 3 Structure of this study

2 Literature

Firstly, this chapter explores the approaches of demand management strategy. The characteristics of demand management strategy approaches and its ways to achieve effectiveness in services are reviewed in order to identify the most suitable demand management strategy to reduce the amount of insignificant testing and create more effective services. Secondly, to determine what characteristics effective healthcare services possess, the metrics of effectiveness within healthcare services is explored through prior research. Metrics varies depending on the type of healthcare service, thus the specific metrics for diagnostic testing will be discussed in this chapter. Lastly, due the objective of this research and to identify ineffectiveness within focal organization, concepts and criteria of an insignificant diagnostic test are discussed and presented based on previous research. In addition, previous similar studies are explored to provide a framework for this study and to identify the challenges and research gaps within this specific field.

2.1 Demand management strategy

This chapter first discusses different demand management methods. The second sub-chapter focuses on exploring the effect of financial incentives on physician's test ordering behavior to provide insight why test ordering is difficult to manage and to affect in welfare countries like Finland. Lastly, this chapter explores the outcomes of demand management strategy.

2.1.1 Approaches to demand management

In this sub-chapter the demand management methods including communication, education and feedback, test request form redesigning, reflex and reflecting testing and guidelines and protocols are discussed. In order to build more effective services, service producers need demand management approaches to focus resources towards patients with the greatest need for it. Demand management strategies have been studied mainly in the field of general laboratory (Fryer & Smellie 2013; Janssens 2010; Lang 2013; Lee et al. 2013). The present study explores and applies the demand management approaches discussed in literature to the field of advanced diagnostic testing such as ENMG testing.

Demand management approaches affect various parts of the test ordering process presented in chapter (Figure 2). One approach is to affect already referring physician's clinical examination and interview procedures by giving feedback or educating physicians (Thomas et al. 2006; Miyakis et al. 2006; Fryer & Smellie 2013; Janssens 2010), so that a referring physician is more critical while ordering diagnostic tests. Another approach is to improve the quality of test referrals. Quality of test requests is improved by redesigning the request form (Bailey et al. 2005; Fryer & Hanna 2009; Fryer & Smellie 2013; Janssens 2010), which then allows referring physician to make more appropriate test referrals according to new request form. All in all, as the present study manages to identify the characteristics and the number of insignificant tests, with demand management approaches such tests can be reduced. This results in treatment plan decision-making and diagnosing of a patient to become more accurate and easier as through more effective services larger number of diagnostic tests will provide value to referring physicians as these tests are more significant and carefully requested.

Dialogue between the key stakeholders – test orderers and test producers – is considered extremely important, as the dialogue enables test producer to educate and give feedback to test orderer, which encourages more significant tests to be ordered and produced (Fryer & Smellie 2013; Janssens 2010; Lang 2013). Especially Lang (2013) claims that real-time notification of inappropriate test request should be sent to the referring physician immediately, so that the physician can either reform the request or alternatively decide not to have the test.

Furthermore, educating the referring physicians is one of the most used and effective demand management strategies (Fryer & Smellie 2013; Janssens 2010; Lang 2013; Lee et al. 2013), although controversial results of education's effectiveness has also been presented (Solomon et al. 1998). Solomon et al. (1998) claim that educating referring physicians about test requesting as well as rewarding physicians for improved requesting behavior only has short-term effects on requesting behavior. One of the reasons why educating physicians is ineffective is that it will only have a local effect (Fryer & Smellie 2013), and educating all referring physicians requires a lot of resources. Due to this it would be ideal that physicians learn adequate test requesting principles already in medical school (Janssens 2010). Fryer & Smellie (2013) present effective educational interventions such as face-to-face sessions, guideline education, information technology, linked information and data feedback which have either reduced laboratory test volumes, costs or insignificance.

On the other hand, today, patients can be demanding and they have more information about their condition than before the internet era, thus patients may no longer rely on physicians' evaluation, but demand for more tests to be performed (Thom et al. 2002). Demanding is related to the level of trust of patient-physician relationship. Low level of trust increased patient's dissatisfaction, and therefore causing patient to demand for more healthcare services such as diagnostic tests (Thom et al. 2002; Bertakis & Azari 2011). Consequently, teaching patients and forming more stable patient-physician relationships within primary care are possible approaches to reduce inappropriate testing even though it is challenging and the effects are not substantial (Fryer & Smellie 2013; Bertakis & Azari 2011). Accepting patients' every wish, intuitively, increases the number of insignificant tests (Lysdahl & Hofmann 2009). In contrast, when patients learn about the consequences of tests and treatments they may even cancel the test, because being afraid that it hurts (Kernick 1998). Reducing pain and discomfort caused to patients unnecessarily, is a valid goal of demand management. Patient education could then be a double-edged sword: it could produce the unwanted effect of having patients refuse ENMG testing because of fear of pain.

In addition to education, giving feedback is considered a useful tool for managing demand (Thomas, Croal & Ramsay, 2006). In order to change physicians' ordering behavior, one should translate feedback into practice (Solomon, Hashimoto, Daltroy, & Liang, 1998). The feedback information could then consist of costs and volumes of the tests a certain physician has ordered, plus other easily and cheaply acquired information. Disadvantages of giving feedback is that forming individual feedback for individual physicians requires a lot of resources (Janssens 2010). Furthermore, while feedback is given to individual physician, it requires stability from the referring quarter in order to keep the learning within the unit and develop it constantly. The feedback information of costs and its effect on physician's requesting behavior is controversial and explored in more detail in Chapter 2.1.2. Local physician turnover may be high, as physicians tend to conduct parts of their training in varying locations, which makes the feedback giving rather challenging. Nonetheless, education and feedback has been shown to be inefficient when performed alone, and Miyakis et al. (2006) showed that only the proper combination of both of these two approaches is effective in managing demand. Furthermore, due to the cost and resource pressure, IT systems including computers and patient information systems tend to be the focus of cost savings (Anderson et al. 2006), thus these systems are not updated constantly to be the latest version. Due to this, IT systems are not agile enough to give easily formulated and accessible feedback to referring physicians, which does not improve communication between stakeholders.

In addition to feedback, redesigning request forms is widely listed as an effective demand management strategy approach according to literature (Bairstow et al. 2006; Solomon et al. 1998; Janssens 2010; Fryer & Smellie 2013; National Coalition of Public Pathology 2012; Shalev et al. 2009; Bailey et al. 2005). Solomon et al. (1998) claim request form design to be an effective approach, which provokes change in physician's test ordering behavior. The greatest potential in request form is achieved, if the request is electronic, giving it the ability to provide information in real-time about retest intervals, or notify referring physician if the request is missing information of patient's symptoms (Fryer & Smellie, 2013). According to Bairstow et al. (2006), electronic test request form maximizes the appropriateness of referrals. In order to succeed with the referral reduction with a new request form design, one must do it carefully; if the request form is too broad and does not make referring physician to ask the right questions from patient concerning the assumed diagnosis of patient, it will not have the effect of reducing the referrals and does not function as a checklist for the physicians (Janssens 2010).

In addition to request form redesigning, including a *patient's pain drawing* to the referral would help with the evaluation of test appropriateness. Studies have shown that the *patient's pain drawing* helps physician with the evaluation of pain of a patient, since pain is a subjective measure (Felix et al. 2010; Sanders et al. 2006; Spyridonis et al. 2013). Such chart includes a figure of human body to which patient colors his pain, numbness and weakness. This method is widely used in Europe to evaluate the level of pain and what the pain indicates, because only spoken description of pain is inaccurate and may mislead physician with diagnosis formulation. (Spyridonis et al. 2013.)

As mentioned, the electronic request form may inform physician about appropriate retest intervals. Notifying the referring physician about the tests that are requested too often could reduce the number of tests done (Fryer & Smellie 2013; National Coalition of Public Pathology 2012). Unnecessary control tests are difficult to identify as the minimum retest interval varies depending on the patient's clinical status (Janssens 2010). Effective minimum retest intervals have been studied in the field of general laboratory, such as the research of effective retest interval for HbA1c testing for monitoring the Type I and Type II diabetes (Driskell et al. 2012). While retests and unnecessary controls are easier to implement in the general laboratory setting and in their request forms, the optimal retest interval is difficult to determine for ENMG tests as it highly depends on the patient's clinical status. The control test interval for ENMG test is advised to be approx. 2-4 months as if neurophysiologist recommends it (Heikkilä & Oja 2015). Nonetheless, control test interval for ENMG test varies depending on the patient's symptoms and disease (ibid.).

Reflective and reflex testing is used widely for laboratory test demand management. These two approaches provide effective and efficient diagnoses to be made. In reflective testing, the test-performing physician adds tests to an order without consulting the referring physician, based only on the clinical question asked. In reflex testing, the test-performing physician automatically performs additional more specific and expensive tests to patient if the first performed cheap and explorative test gives an abnormal result. (Srivastava et al. 2010). According to Janssens (2010), reflective and reflex testing prevents useless screening with expensive tests, but it requires that the clinical examination is done properly before referring to any diagnostic tests. Reflective testing is actually an approach of demand management neurophysiologists want to avoid, as they do not want to perform any additional expensive tests to patient, rather they prefer to have patients, who are in advance meant to be examined with especially ENMG test. This reduces the time spent per patient, shifts responsibility away from neurophysiologist, and provides resource savings. On contrary, neurophysiologists tend to use the reflex testing, as they begin with the less extensive cheaper ENMG test and then if the result is abnormal, they conduct more extensive ENMG test. In other words, neurophysiologists aim to focus ENMG test specifically as possible, and if the cause is not found, then more extensive search with ENMG is required. ENMG test itself is already a very specific and expensive test, thus conducting more common and explorative tests before referring to ENMG is advisable. According to Finnish national Current Care Guidelines, referring i.e. a carpal tunnel syndrome patient to operational treatment based only on ENMG test is not recommended, but the referring physician needs to have other evidence to back ENMG test results before the patient is operated (Suomalainen Lääkäriseura Duodecim 2013).

Besides reflective and reflex testing, different test combinations and packages are offered as one test product to referring quarters, especially in laboratories. These test packages are predefined and usually these are tailor-made to test the function of a specific organ. (Fryer & Smellie 2013; Janssens 2010.) However, Fryer & Smellie (2013) claim that test packages should be made according to diagnosis, thus it would enable more efficient and effective testing if the requestor defines which diagnosis he/she is looking for with the tests. Albeit test packages might be seen as threat of misdiagnosis, according to demand management literature review by Solomon et al. (1998) they may also remind the referring physician to order the right tests given that they have an expected diagnosis, and thus reduce the number of superfluous requesting. Like test request forms, predefined test packages should not be overly restrictive, as then referring physicians tend to order additional extra tests to the package which increases the number of inappropriate tests (van Wijk et al. 2001; Janssens 2010).

Although guideline and protocol development is presented as one of the most effective demand management strategy approaches (Solomon et al. 1998), it is regarded to be very slow and burdensome, as formulating and communicating a new guideline or protocol in medicine requires years of research and commitment (Janssens 2010). However, Warren et al. (2010) succeeded to reduce laboratory test demand by 20% by reformulating the clinical protocol. If redesigning the guidelines and clinical protocols is not possible with the limited resources, the dialogue between the test producer and test requestor could be a good enough method to achieve the same outcome (Janssens 2010).

Besides clinical protocols, some test referrals require acceptance by a diagnostic test specialist before the test is performed. This approach of demand management is called test request vetting. Fryer & Smellie, (2013) stated that urine toxicology screening test request vetting reduced the number of insignificant tests and resulted in cost savings of 30 000 pounds in general laboratory (Figure 4). According to Mondelli et al. (1998), ENMG referrals from a neurosurgeon, neurologist, orthopedist, rheumatologist or physiatrist should be accepted directly by the CN laboratory, and referrals by other physicians such as primary care physicians should be evaluated by neurologist before sending a patient to ENMG test to CN laboratory. Mondelli et al. (1998) claimed that this approach would enable reduction in electromyography (EMG) requests by 25 % in Siena, Italy at the time of the study. Electromyography (EMG) tests examine the function of muscles, while electroneurography (ENG) examines the function of nerves also referred to as nerve conduction study (NCS).

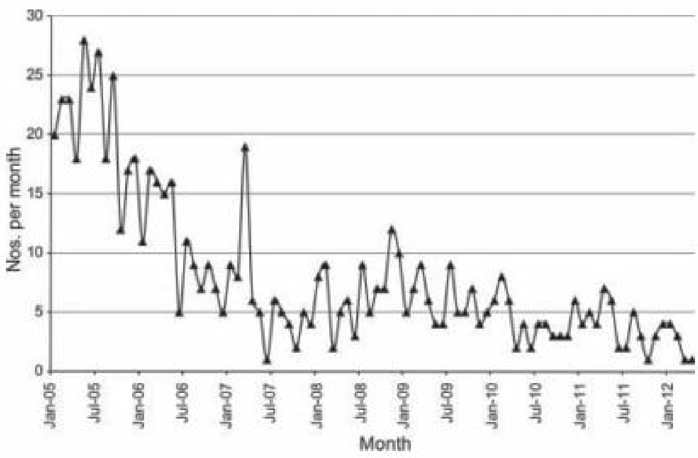

Figure 4 Figure adapted from Fryer & Smellie (2013, p.64). Implementing test request vetting for urine toxicology screening tests in June 2006, resulting in reduced request volume and cost savings of approx. 30 000 pounds per year.

According to literature, the costs and charges of tests has been shown to affect physicians' test ordering (Gama et al. 1992; Seguin et al. 2002). In contrast, Solomon et al. (1998) presented that providing cost feedback alone is not as effective as combining different demand management methods. Furthermore, Hoey et al. (1982) presented that many tests are requested due to certain patient condition reasons, and price information does not affect these kinds of tests. The effect of financial incentives on test ordering behavior is examined in more detail in Chapter 2.1.2.

2.1.2 The effect of diagnostic tests' charges on physician's test ordering behavior

The financial asymmetry between healthcare service orderer and producer causes a part of ineffectiveness within welfare countries' healthcare system, because the referring physician has no personal cost pressure or performance-related pay affected while ordering and purchasing healthcare services for patients. Due to the interesting phenomenon of financial asymmetry, the effect of financial incentives is discussed in this chapter. The outcome and result from financial incentive driven healthcare system is discussed in more detail in chapter 2.1.3.

The feedback of charges of diagnostic tests given to referring physicians reduces the number of diagnostic laboratory tests ordered (Tierney et al. 1990; Seguin et al. 2002; Miyakis et al. 2006; Gama et al. 1992; Sood et al. 2007; Cummings et al. 1982), thus balances the financial asymmetry. The feedback of the costs of diagnostic laboratory tests was usually given to the physicians via the request form (Seguin et al. 2002;Cummings et al. 1982). However, according to Seguin et al. (2002) the most expensive and unnecessary diagnostic laboratory tests reduced, but the volume of some tests stayed the same. These tests were seemingly considered absolutely essential for the patient diagnosis. The cost information did not affect the volume of clinically absolute diagnostic tests, thus some tests should still be considered as price insensitive tests.(Seguin et al. 2002.) The same trend was found in a research done in a cardiac care unit by Warren et al. (2010). Both of the aforementioned studies examined the volumes of common laboratory tests, which might have been more price sensitive, unnecessary tests.

However, electrodiagnostic tests are more specialized tests, and thus might be seen more price insensitive. On the other hand, Seguin et al. (2002) came to the conclusion that the volume reduction occurred especially among the most expensive tests, which trend might also function among the more expensive electrodiagnostic tests like ENMG test. All things considered, physicians' test ordering was affected by the cost information (Tierney et al. 1990; Seguin et al. 2002; Miyakis et al. 2006; Gama et al. 1992; Sood et al. 2007; Cummings et al. 1982). On the other hand, physicians act on the behalf of a patient (Eisenberg 2002), thus physicians are expected to order all the tests required to diagnose the patient with high quality. This expectation and the role of physicians result in price insensitivity, which may be one reason for the difficulty of reducing the number of insignificant tests.

Although available cost information affects physician's test ordering, physicians were often unaware of the costs of diagnostic tests even though the tests were commonly used (Miyakis et al. 2006; Seguin et al. 2002; Allan & Lexchin 2008). The lack of cost information was especially noticed among trainee physicians (Miyakis et al. 2006), who even tend to order more tests than seniors (Winkens et al. 1995). Reasons for this could include knowledge on modern test technologies and services, or weaker self confidence and thus increased reliance on diagnostic tests (Winkens et al. 1995). On the other hand, Miyakis et al. (2006) noticed that the most expensive test in their research was used more appropriately than the other tests even though physicians were unaware of the charges of these tests. As a result, it seems that cost unawareness does not cause the inappropriate use of tests (Miyakis et al. 2006). However, these studies were made with common and mostly price sensitive laboratory tests, thus physicians ordering behaviors with more expensive and specialized electrodiagnostic tests may be different from that with common laboratory tests. In other words, there is not enough knowledge concerning price insensitive expensive tests to predict the effect of cost on electrodiagnostic test requesting.

Furthermore, researchers were uncertain of sustaining the cost reductions in the long run, as previewing the charges in request form would reduce the costs to some extent but it would be difficult to maintain constant cost reduction over time (Seguin et al. 2002; Warren et al. 2010). In addition, providing feedback and education to physicians is not cheap, because both effective education and feedback require face-to-face sessions, and producing and communicating the information to physicians is time and resource consuming (Eisenberg 2002). Schroeder et al. (1984) calculated the cost of education and feedback and concluded that the cost of intervention consumed the savings caused by the intervention. Controversially, Gama et al. (1992) reported that the cost of feedback in their research was £530 while the cost savings was £3027 calculated in revenue expenditure, showing savings in total costs even after the investment of feedback production (Gama et al. 1992).

2.1.3 The outcome of demand management

This chapter discusses the outcomes achieved with demand management approaches. In addition to results from demand management strategy usage, the challenges of demand management are explored.

The consequences of reducing demand are assumed to be misdiagnosing patients, resulting in more costs as patients are undertreated and referred more often to secondary care. Interestingly, however, some studies have found out that reducing unnecessary and inappropriate tests does not affect patient outcome, which is measured by length of stay, mortality and readmissions to hospitals (Warren et al. 2010; Winkens et al. 1995; Gama et al. 1992; Seguin et al. 2002). According to Winkens et al. (1995) reducing the number of unnecessary tests reduced the number of specialized referrals without affecting the patient outcome. Due to this phenomenon, display of the charges of test on request forms for referring physicians to see, redesign of request forms, and education of referring physicians could actually provide great savings with similar quality of care.

Demand management helps to match the demand to supply by either cutting inappropriate demand, increasing preventive demand, or coping better with existing demand through effective care (Figure 5) (Pencheon 1998). Demand management has various ways to reduce the number of insignificant diagnostic tests (Lee et al. 2013; Janssens 2010; Fryer & Smellie 2013; Bernardy et al. 2009), which results in shorter waiting lists. Then limited resources are used more efficiently, and thus costs are reduced. As a consequence better patient clinical outcomes are achieved (Winkens et al. 1995), which may be due to the waiting time reduction which then enables patients with actual need for the service to get the help they need on time. Both the healthcare professionals and patients form the demand (Pencheon 1998), thus it is essential to affect both parties in successful demand management.

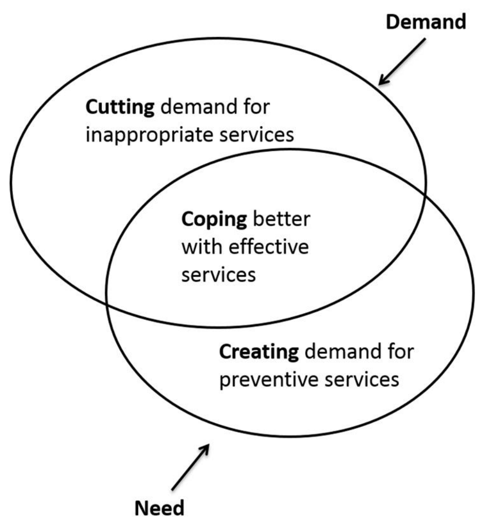

Figure 5 Framework for the management of the demand for healthcare services, adapted from (Pencheon 1998)

Demand management has an impact on laboratory, patient and clinical outcome (Fryer & Smellie 2013). Laboratory impact is measured in number of tests and cost, while patient and clinical impact are more difficult to measure. Patient impact shows i.e. in the number of false positive and negative results, discomfort and anxiety and in the number of days off work (Janssens 2010). Furthermore, the effectiveness of demand management is important to measure in the light of clinical and patient outcomes besides laboratory impact. Demand management outcomes should not be limited to the reduction of test volumes and costs, only we must see the overall picture, and how reduction of inappropriate test affects patient outcome. As over-testing is reduced and under-testing is noticed, the result will be fewer referrals to secondary care (Winkens et al. 1995). Furthermore, it will result in fewer hospital stays and better quality of patient life due to fewer false positive test results and less distress of testing. The values and benefits achieved with more effective healthcare services are discussed in more detail in chapter 2.2.2.

Solomon et al. (1998) divided the demand management strategies into three different categories: predisposing factors, enabling factors and reinforcing factors. Predisposing factors include education and training, which are meant to help physicians notice the need for change. Enabling factors include change of the test request protocols, agreement on new guidelines and redesign of the request form. These factors enable physician to act according to new guidelines and to achieve the change in requesting behavior. Reinforcing factors give feedback and engage physicians to keep track of the costs of diagnostic test, which encourages physicians to keep the change in their behavior and rewards physicians when they succeed. Nevertheless, the enabling factors were considered the most powerful in demand management (Solomon et al. 1998). On the other hand, successful results by using demand management are achieved with combining these different factors according to other prior studies (Miyakis et al. 2006; Solomon et al. 1998; Fryer & Smellie 2013; Fryer & Hanna 2009).

Laboratory test volume reductions vary between 7.9% - 79% depending on the demand management strategy used (Bunting & Van Walraven 2004; Gama et al. 1992; Bailey et al. 2005; Seguin et al. 2002; Shalev et al. 2009). A study done by Bunting & Van Walraven (2004) reported that a combination of multifaceted education and feedback reduced their test volumes by 7.9%. Gama et al. (1992) achieved a test volume reduction of 27% with feedback strategy after the first year intervention. The cost feedback given to physicians reduced the cost expenditure by 22% within 2 months (Seguin et al. 2002). Bailey et al. (2005) redesigned a request form for laboratory tests by reducing the preprinted tests on the request form and achieved a test volume reduction of 27–79% depending on the laboratory test over a two-year period. Similar results were reported by Shalev et al. (2009) in which the test volume increased by 60–70% with tests which were added to the request form and reduced the excluded tests by 19.2–27 %. All in all, demand management approaches reduce test volumes and costs, enabling shorter queues to diagnostic tests. However, all of these studies examined the common laboratory tests, which might differ from more specialized and expensive electrodiagnostic tests in characteristics. Thus the same approaches, implemented for electrodiagnostic tests, may give different results.

Yet, many studies observed that only few demand management approaches sustained a long-term change after the invention (Warren et al. 2010; Seguin et al. 2002; Miyakis et al. 2006; Gama et al. 1992; Eisenberg 2002; National Coalition of Public Pathology 2012). The combination of education and feedback was considered an effective demand management approach, but the change in test volumes did not last long (Eisenberg 2002; Miyakis et al. 2006). Due to the short-term effect and the expensive costs caused by providing the feedback and education, education and feedback giving could be considered as not efficient enough demand management approach. Controversially, Gama et al. (1992) reported that individual feedback generated a sustained the change. Furthermore, Winkens & Dinant (2002) claim that feedback is an effective demand management approach, when it is possible to translate the feedback information to concrete practice and it gives a possibility to physicians to participate in the process.

Nonetheless, the greatest challenge after test volume and cost reduction is to sustain the change (Warren et al. 2010; Seguin et al. 2002; Miyakis et al. 2006; Gama et al. 1992; Eisenberg 2002). The success and sustainability of the demand management intervention depends on many actions other than the intervention itself (Solomon et al. 1998; Winkens & Dinant 2002). These additional actions that sustain the change of an intervention include use of multiple demand management approaches, engagement of the clinical staff and especially the senior level, ensured support from senior and lead clinicians, and attempt to keep approaches simple to be easily adopted into practice. (National Coalition of Public Pathology 2012; Solomon et al. 1998).

2.2 The effectiveness of healthcare systems

In this chapter the concept of effective healthcare services and how to measure effectiveness are examined. As presented in Chapter 2.1 demand management is a tool especially for healthcare professionals to encourage patients and other professionals to use healthcare services more appropriately. Furthermore, there should be tools for healthcare service consumers. Such a tool would enable them to base their decision on relevant information, which tells how and what to consume. One type of such information is to present the actual effectiveness of services to consumers, patients and political decision makers, which then could encourage the whole healthcare system to become more effective. However, in order to show results from effective service, one has to determine and decide how to measure effectiveness. Therefore also the metrics of effective services are explored.

The first effectiveness studies emerged during 1960s and 1970s, and they calculated costs of diseases (Johannesson & Jönsson 1991). Effectiveness studies compare the input and output of a health care intervention, which can be either a diagnostic test or treatment. The output consists of health benefit and possible harm. The input is mostly measured in monetary terms, such as workforce required or the cost of a new technology. (Kernick 1998.) The need for effectiveness studies originate from increased demand and healthcare costs as well as from limited resources of healthcare (Chalkidou et al. 2009; Kaplan & Porter 2011). More efficient resource allocation makes cost savings possible (Eichler et al. 2004).

This trend is nowadays evident in many healthcare sectors in numerous western countries both in diagnostic testing and in common physician appointments.

One of the first approaches used to evaluate effectiveness was the human capital approach. Human capital approach is based on a method of calculating the benefit of a treatment or technology as reduced costs and reduced working resources. Due to the several limitations of the human capital approach, such as it lacked the quality of life in the evaluation and did not take into account workers outside permanent workforce, the cost-effectiveness analysis (CEA) was developed. Furthermore, the human capital approach calculated disease-specific costs, which did not relate to any treatment or technology, thus the approach did not provide anything to decide on. The most important advantage of the CEA compared to the human capital approach is that it takes into account the number of gained life-years. (Johannesson & Jönsson 1991.) In CEA, the input measured in money is compared to outcome which is commonly measured in mortality, morbidity and life-years gained (Johannesson & Jönsson 1991; Singer & Applegate 2001; Kernick 1998). However, CEA also has limitations concerning its comparability, as the calculations vary in terms of timeframe used, sorts of costs included, and type of life-years gained (Johannesson & Jönsson 1991; Singer & Applegate 2001; Kernick 1998; Kaplan & Porter 2011; Neumann et al. 2010). With CEA, decision makers are able to define the absolute benefit gained through a new health intervention, but it is difficult to compare the results between different interventions because of the varying calculation methods (Johannesson & Jönsson 1991; Kernick 1998).

Besides cost-effectiveness analysis, also cost-utility and cost-benefit analyses were developed (Johannesson & Jönsson 1991; Singer & Applegate 2001; Kernick 1998). These methods have emerged from CEA, and are comparable to CEA. Both compare the output of the intervention to the input, but cost-utility and cost-benefit differ in term of how the output is measured. In the cost-utility analysis, the effectiveness is measured usually with quality adjusted life years (QALYs) (Kernick 1998; Singer & Applegate 2001). The idea in cost-utility analysis is to maximize the number of quality adjusted life years with the given healthcare budget. Like in CEA, comparability between two studied treatment methods or diagnostic tools is weak due to varying types of costs and life-years used in the calculations. (Johannesson & Jönsson 1991.) In cost-benefit analysis, both inputs and outputs are measured in economic units, thus the comparison of inputs and outputs is made easier and generally more acceptable. On the other hand, cost-benefit analysis has a drawback because it values the health benefit in financial terms, which is not simple to do (Johannesson & Jönsson 1991; Singer & Applegate 2001). Conversely, Kaplan & Porter (2011) claim, that the difficulty of calculating the health outcome benefit in monetary units is only a myth, behind which healthcare providers hide in order to ensure high reimbursement pays from government and insurance companies.

As the cost-benefit and cost-utility analyses did not take into account the opinions of health intervention participants, the contingent value method (CV) was developed. In the CV method, the willingness to pay and willingness to accept is studied with a survey from healthcare consumers and decision makers.

(Johannesson & Jönsson 1991.) The willingness to pay is measured as the maximum amount of money a person is ready to yield to secure that the proposed treatment or intervention is performed (Johannesson & Jönsson 1991; Eichler et al. 2004). In contrast, willingness to accept is the minimum amount of money a person accepts to take if the treatment or intervention is not produced. (Johannesson & Jönsson 1991.)

Comparative effectiveness research (CER) is a broader concept, encompassing the use of the aforementioned analysis methods (Goehler & Gazelle 2014; Gazelle et al. 2011). The aim of CER is to compare the current standard of patient care to alternatives, such as new treatments brought about by new technologies (Chalkidou et al. 2009; Goehler & Gazelle 2014; Gazelle et al. 2011). These comparisons are then conducted with the help of e.g. CEA, cost-benefit analysis or cost-utility analysis (Goehler & Gazelle 2014; Gazelle et al. 2011).

Although effectiveness research is highly praised and many healthcare decision makers encourage everybody to use it (Kaplan & Porter 2011; Neumann & Tunis 2010), there are controversial perspectives to effectiveness evaluations. These limitations of effectiveness research include the varying ways of calculating both the inputs and the final outputs (Kaplan & Porter 2011). Furthermore, a lot of debate has been about utility weighting while evaluating the outputs. In fact, the Centers for Medicare and Medicaid Services (CMS) in the United States have explicitly excluded cost-effectiveness results from the coverage of decision-making due to the danger of favoring only specific population groups as a result of a certain way of calculating the cost-effectiveness. (Neumann & Tunis 2010.) There has also been discussion of whether effectiveness calculations emphasize too much certain diseases and ages and put patients into unequal positions (Neumann et al. 2010). On the other hand, effectiveness studies should evaluate the output compared to input, making use of utility functions: e.g. a life-year of a working-age patient is more expensive than a life-year of 80-year old patient, even though this kind of thinking may be politically incorrect. Neumann et al. (2010) found that the worry towards the metric of quality adjusted life years (QALYs) was about if it emphasizes unequally the importance of younger and healthier than older generations' life-years as younger have more high quality life-years to gain Metrics of effective healthcare service

2.2.1 Measuring effectiveness within healthcare

In order to improve the effectiveness of health care services, metrics are needed to measure and analyze the development. In this chapter the most common metrics used for effectiveness evaluation within health care services is explored. In addition, also metrics most suitable for analyzing effectiveness in diagnostic testing are discussed.

Effectiveness of healthcare services is often measured by patient outcome, such as patient's functional status, quality of life, disability, clinical events, mortality or morbidity (Gazelle et al. 2011; Seguin et al. 2002; Singer & Applegate 2001; Chalkidou et al. 2009). The most common effectiveness metric is the quality of life-years gained, which is often compared to costs, thus measured by cost per quality-adjusted life year (cost per QALY). In QALY, health utility is determined according to preference weighting. The weighting factor is calculated to each state of health, where the value of zero corresponds to death and value of one to full health. The QALY is then the product of the health utility score and the time spent in that state of health (Singer & Applegate 2001; Kernick 1998; Räsänen et al. 2006). Limitations of QALYs include the difficulty in evaluation, and especially, equal evaluation of the quality of a life-year with a specific disease (Räsänen et al. 2006). As a result, the comparison of QALYs from different cost-utility analyses is problematic.

QALY, mortality, patient's functional status and morbidity are difficult to measure as the outcome of a diagnostic test (Drummond et al. 2009; Lee et al. 2010). One of the effectiveness metrics used to measure diagnostic testing in terms of health benefit is *medical value*. *Medical value* is the value of information gathered from diagnostic test and the value of correct diagnosing, which affects the final patient outcome as well. (Lee et al. 2010; Ferrante di Ruffano et al. 2012.) Another effectiveness metric of diagnostic testing used is *the planning value* to the patient. This refers to patient's ability to adjust life according to the test result. The third effectiveness metric relevant to patient is *the psychic value*, which refers to patient's mood and life changing according to test results. (Lee et al. 2010; Thornbury et al. 1999.) "*The value of knowing*" has also been emphasized in the past years as important non-medical value of diagnostic testing. It represents the psychic value of diagnostics to patient, as knowledge on own health condition is important to patient's psychological well-being. (Lee et al. 2010.) *Planning value*, *psychic value* and the '*value of knowing*' all refer to the physiological effects of a diagnostic test to a patient. In addition, patients feel anxiety and distress before the test if they are unaware of the test procedure (Thornbury et al. 1999), especially in ENMG testing (Kothari et al. 1995). This also affects their psychological well-being, thus it can be considered as a *psychic value* to a patient to know about the test beforehand (Lee et al. 2010). These metrics are assessed in a method called *patient utility assessment* presented by Fryback & Thornbury (1991). Usually patients value diagnostic testing highly due to the importance of "*value of knowing*", which may cause patients to demand certain tests in order to get answers concerning their symptoms.

The value and benefit measured in effectiveness depends on the role of the participant in a diagnosis situation (Gazelle et al. 2011; Lee et al. 2010). *Medical value* is the most important value to physician, who diagnoses and decides about patient management according to the results (Gazelle et al. 2011; Thornbury et al. 1999; Lee et al. 2010). Furthermore, the payer is usually mostly concerned about *medical value* in the context of costs and overall patient outcome. At the same time, an important factor for the patients is the psychological value of a test. (Lee et al. 2010.)

Nonetheless, diagnostic tests are difficult to evaluate from the patient outcome perspective as diagnostic tests primarily affect the physician who sets the diagnosis and decides the treatment plan (Lee et al. 2010; Ferrante di Ruffano et al. 2012). However, this effect on treatment plan decision-making in turn affects patient outcome eventually after the decision and treatment is performed (Fryback & Thornbury

1991; John W. Loop & Lusted 1978; Kaplan & Porter 2011; Ferrante di Ruffano et al. 2009; Bossuyt & McCaffery 2009). Due to the problematic connection between diagnostic test and final patient outcome, the concept of *diagnostic thinking efficacy* was introduced by Fryback & Thornbury (1991). *Diagnostic thinking efficacy* evaluates how physicians' decision-making about diagnosis is affected by a diagnostic test result. In addition, Fryback & Thornbury (1991) presented the *therapeutic efficacy*, which evaluates the effect of a test result on a physician's treatment plan decision-making. Fryback & Thornbury (1991) presented a framework for evaluating the efficacy of a diagnostic test. The framework (Figure 6) includes six levels: technical, diagnostic accuracy, diagnostic thinking, therapeutic, patient outcome and social efficacy. Of these methods, the technical and diagnostic accuracy efficacy methods are appropriate for measuring new technology and how it effects patient's treatment. In contrast, the patient outcome and social efficacy methods are time-consuming to evaluate, and these measure the longer term effects of diagnostic testing. (Thornbury et al. 1999.) Hence, in this study we focus on evaluating the effectiveness of ENMG test from the diagnostic thinking and therapeutic efficacy point of view.

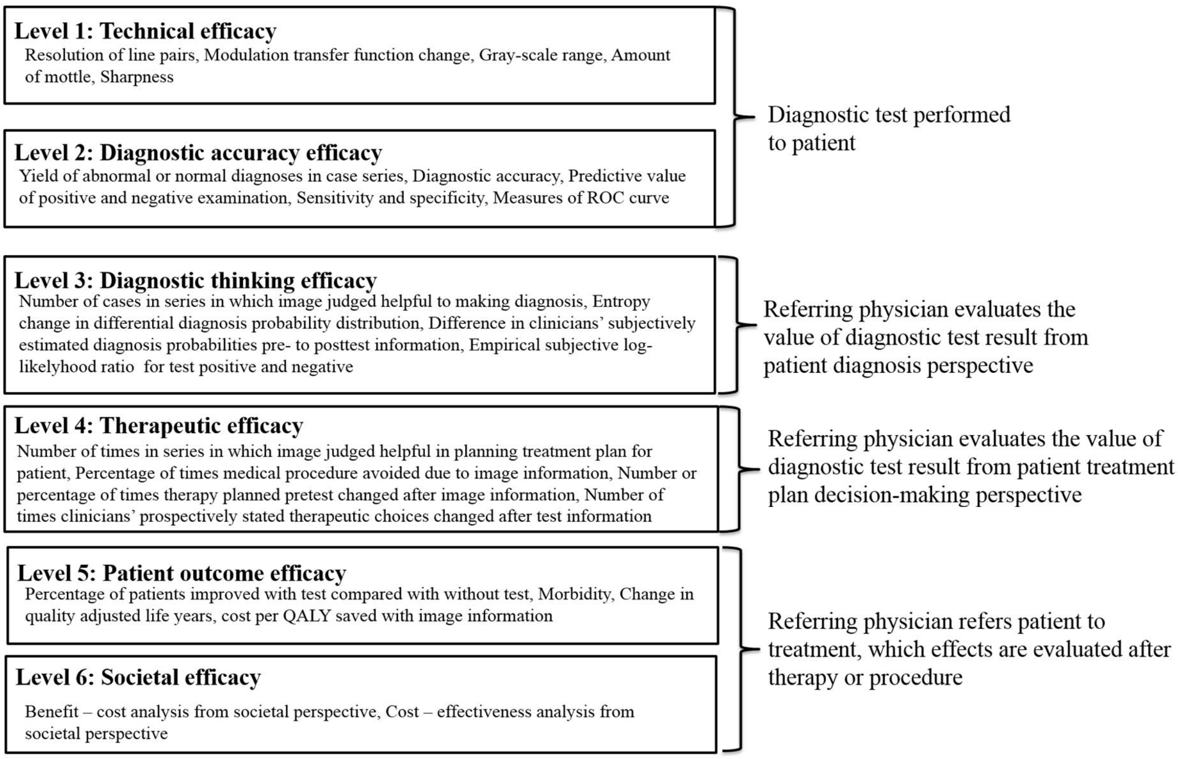

Figure 6 Hierarchical model of the efficacy of diagnostic tests, adapted from Fryback & Thornbury (1991)

As the purpose of a diagnostic test is to inform a physician about a patient's condition, the most valuable impact of a diagnostic test is how it affects a physician's decision and treatment plan decision-making (Janssens 2010). On the other hand, Bossuyt et al. (2012) argued that it is not relevant whether a test provides information, but rather, whether it provides useful information or whether it is used correctly and efficiently, which determines whether a diagnostic test has been effective from the clinical care perspective. As diagnostic thinking and therapeutic efficacy are important metrics in diagnostic testing, many studies have used either surveys done to referring physicians questioning the impact of a diagnostic test on diagnosing or compared the pre and post diagnoses in order to determine whether a test was useful to physician's decision-making process (Shepherd 2010; Komur et al. 2014; Rigler & Podnar 2007).

On the other hand, Ferrante di Ruffano et al. (2012) presented that the impact of a diagnostic test should not be removed from the final patient outcome and especially not from the influence of other tests, thus the impact of diagnostic test effectiveness should be evaluated from the final patient outcome perspective. This influence of diagnostic test on patient outcome is shown in Figure 7 (Ferrante di Ruffano et al. 2012). The effect of test result on diagnosing and treatment plan decision-making is considered as a short-term patient outcome metric, while mortality, procedure side effects and morbidity are considered to be long-term patient outcomes (Thornbury et al. 1999). The present study does not remove the impact of diagnostic testing from patient outcome, but it only evaluates the effectiveness from the short-term patient outcome perspective, the diagnostic yield and therapeutic yield are shown in Figure 7. The present study uses short-term effects as effectiveness metric, similarly as Shepherd (2010) evaluated the effectiveness of EMG tests by the number of abnormal results, the number of diagnoses an ENMG test confirmed, whether the EMG test changed the treatment plan, and whether the ENMG test provided new diagnosis alternatives. In other words, Shepherd (2010) also used diagnostic thinking and therapeutic efficacies to determine the effectiveness of EMG test.

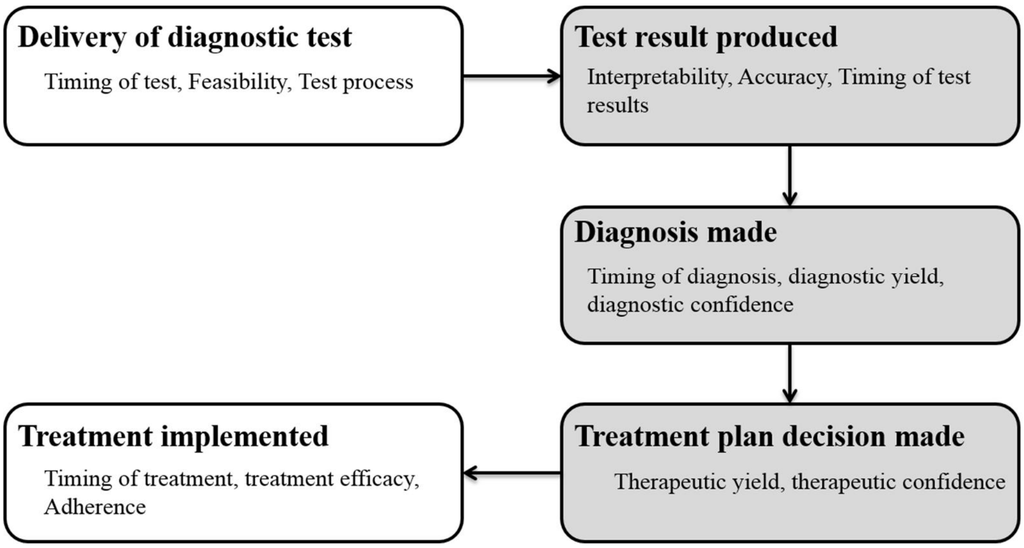

Figure 7 Attributes of test-treatment pathway that affect patient health, adapted from (Ferrante di Ruffano et al. 2012)

The concepts of *utility* and *value* of a diagnostic test are commonly used to describe the same concept, although *utility* and *value* mean different things (Singer & Applegate 2001). *Utility* is a subjective measure depending on the preferences of the evaluator and it is often used in assessing clinical or patient outcomes such as QALY (Singer & Applegate 2001). *Value*, on the other hand, is more of an absolute measure: medical *value* of a diagnostic test is measured in the ability of the test to provide relevant information to physician, which enables correct diagnosis and treatment of a patient (Lee et al. 2010). Although *value* is considered a more objective metric than *utility*, the methods used to measure *value* are diverse (Johannesson & Jönsson 1991; Singer & Applegate 2001; Kernick 1998) and difficult to perform with proper comparability (Johannesson & Jönsson 1991; Simonen et al. 2012; Neumann & Tunis 2010). As *utility* is considered controversial, so is *value* due to the aforementioned problems.

Clinical utility and *patient utility* measure different preferences (Bossuyt et al. 2012; Thornbury et al. 1999). *Clinical utility* concerns clinical events such as the number of correctly diagnosed patients (Bossuyt et al. 2012), whereas *patient utility* is used for measuring patient's well-being (Thornbury et al. 1999). *Patient utility* assessment includes same metrics as Lee et al.'s (2010) *psychic value* of a diagnostic test: patients anxiety, planning value, expectations of future health and quality of life, and Thornbury et al.'s (1999) metric of patient's willingness to pay. All in all, *patient utility* is evaluated more from the psychological perspective, while the clinical utility takes into account the actual impact of an intervention. In other words, *patient utility* measures the impact on individual, while *clinical utility* measures the impact on population.

As stated in this chapter, the effectiveness analyses are not as comparable as analysis users would hope (Johannesson & Jönsson 1991; Singer & Applegate 2001; Kernick 1998). Thus, diagnostic thinking and therapeutic efficacies are more commonly used in evaluating diagnostic testing, as it does not require cost effectiveness or cost-benefit analyses and utility preferences (Fryback & Thornbury 1991). Furthermore, final patient outcome is difficult to measure for a diagnostic test (Drummond et al. 2009). However, there are contradicting perspectives to whether patient outcome should be separated from the evaluation of effective diagnostic test and evaluate in isolation (Lijmer et al. 2009). Lastly, the reason for healthcare managers' avoidance of effectiveness data is, according to Simonen et al. (2012), the incomparability and reliability of effectiveness metrics such as quality adjusted life years (QALY). To conclude, the healthcare system would need more reliable and agreed effectiveness metrics to evaluate effectiveness.

2.2.2 The outcome of effective healthcare services

This chapter discusses the outcomes achieved through more effective healthcare services. Furthermore, the financial incentives are discussed, in order to explore economically oriented environment a healthcare system would actually require becoming more effective. Ineffective service production exists due to the large amount of demand, which is created by patients with both actual need and no need for the service. Demand is not prioritized with the thought of how much the service actually increases the medical value. Besides demand management strategies, which help service producers to guide referring physicians, also incentives are needed to change the requesting behavior and to control and prioritize the demand.

The measured outcomes of effectiveness of healthcare service often used are divided into patient outcome, clinical outcome and economical outcome (Bossuyt et al. 2012; Thornbury et al. 1999; Fryback & Thornbury 1991; Petrou & Gray 2011; Gazelle et al. 2011; Ferrante di Ruffano et al. 2012). These well-established outcomes guide the development of a healthcare service as decision makers encourage improving the selected outcomes. Patient outcome is measured as a psychological effect on patient's welfare such as how a patient perceives the health state after an intervention (Lijmer et al. 2009; Lee et al. 2010). Clinical outcome is the absolute success of a clinical event such as the number of tumors detected or blood pressure reduced (Chambers et al. n.d.). Economical outcome reflects monetary results, such as the indirect and direct costs of an intervention vs. the final savings. Often, economical outcome is measured where two or several different interventions are compared in terms of their cost-effectiveness. (Eichler et al. 2004; Chalkidou et al. 2009).

In order to achieve effective healthcare services, the financial asymmetry between healthcare service producer and orderer has to be improved. The financial incentive should be implemented into the mindset of referring physicians, which would then change behavior of healthcare professionals (Flodgren et al. 2011). The financial asymmetry leads to a situation, where physicians are free to order every service their patient does or does not need without it affects physician's economics or performance. Because the ordering and purchasing of services affects the referring quarter i.e. primary care unit as whole, the cost of services does not show to the referring physician, who ordered the service and referred the patient. In addition, Kaplan & Porter (2011) argue that the costs of healthcare services are usually calculated to match with the agreed reimbursement given by government or insurance companies rather than to match the actual resource usage, which creates no incentives to service producers to improve their resource utilization. This also reflects the situation of public sector, where production units order services from one another without showing the actual costs of a service to the service-ordering physician, which would create healthy concern about the resources used. As discussed in Chapter 2.1.2, physicians usually are unaware of the costs of diagnostic tests, even though they order these tests almost on a daily basis.

The British national healthcare system's Quality and Outcomes framework (QOF) is a good example of recent experiment to improve the effectiveness (Roland 2004). Fleetcroft et al. (2012) argue that the British national healthcare system (NHS) did not succeed to maximize the health gains of population with its QOF supported by financial incentives, which incentives depend on certain selected indicators of health. Conversely, Campbell & Lester (2008) state that NHS succeeded to improve the preventive healthcare and the behavior of healthcare professionals by setting indicators for disease-specific processes, which improved the effectiveness of healthcare system as well.

NHS implemented 146 indicators in 2004 to measure effectiveness of their healthcare system. The General Medical Service (GMS) linked up to 20-25 % of the general physician's income to these indicators, which was done in order to change their behavior and improve effectiveness. (Roland 2004; Guthrie et al. 2006.) NHS's idea was to create indicators concerning the health condition of patients with chronic diseases, encouraging physicians with monetary incentives to improve overall health of the whole population, as chronic disease patients place a burden on the healthcare budgets (Marshall & Smith 2003).

What seems to have happened is that they overlooked the possibility of neglecting non-incentivised diseases; the monetary bonuses were only paid for curing chronically ill patients with bonus related indicators (Doran et al. 2011). The scare scenario, where physicians only select optimal patients for selfish economical gains (Mangin & Toop 2007; Roland 2004), eventually proved to be somewhat true (Doran et al. 2011). Due to the possibility of gaming for individual gain and reducing the quality of healthcare, healthcare decision makers should very carefully choose their effectiveness metrics and indicators for physicians. NHS used an effectiveness improvement method of changing physicians' behavior by creating financial incentives. The debate of whether setting financial incentives to physicians to improve the healthcare system in Finland depends on the metrics and incentive system chosen.

All in all, emphasizing and improving effectiveness of healthcare interventions and services is targeted to increase the quality of health outcome with continually reducing resources. In other words, with more effective healthcare we get more benefit with fewer resources. To sum, the idea of effectiveness is to reallocate current and future resources, so that the maximum health benefit is achieved (Räsänen et al. 2006).

2.3 Insignificant diagnostic tests

In this chapter, the characteristics of insignificant diagnostic tests and inappropriate test requests are explored and discussed. In addition, the reasons for ordering such tests are shortly presented based on prior studies. Finally, the outcome of reducing insignificant tests is examined through previous studies.

Prior studies have examined the appropriateness of EMG requests (Cocito et al. 2006; Mondelli et al. 1998; Podnar 2005; Fuglsang-Frederiksen et al. 1995). These studies have shown that 17–25% of the EMG tests conducted were insignificant due to inappropriate EMG test requests. Furthermore, recent studies by Di Fabio et al. (2013) and Mondelli et al. (2014) concerning EMG tests' significance and test referral appropriateness prove this study topic to be timely, as these two studies were published during the last few years and revealed that insignificant EMG tests are performed constantly and that it depends on the quality of test referral and a couple of other factors discussed in next sub-chapter.

2.3.1 Characteristics of an insignificant test

If the effectiveness of a diagnostic test was measured by its effect on physician's diagnostic and treatment plan decision-making (Lee et al. 2010), then an insignificant test could be considered to be a test that does not affect physician's decision-making. However, van Walraven & Naylor (1998) presented such categorization of inappropriateness too implicit due to the fact that diagnostic decision-making never depends on one test only, but rather on several tests. Thus the evaluation of whether a diagnostic test affects decision-making or not, is not enough to determine the significance of a single test. In addition, a diagnostic test that does not affect specifically physician's decision-making may still provide additional and useful information to both patient and physician (ibid.) Similarly, Lijmer et al. (2009) presented that diagnostic tests should not be removed from the context and from the influence of other tests for the evaluation, unlike most of the studies use this method. Furthermore, the setting of evaluating a single isolated test, does not allow experiments; in other words, it is impossible to exactly measure how the decision-making process would have proceeded, if the diagnostic test had not been requested. Moreover, Fuller (2005) states that when considering an ENMG test, the referring physician should estimate the value it provides combined to other already ordered diagnostic tests' results. It would be useful to analyze and evaluate a diagnostic test side by side with other tests. However, since that is difficult, complicated and time-consuming, the impact of a diagnostic test is usually evaluated independently.

However, the definition of an insignificant test depends on the field of medicine. Laboratory tests, radiological test and electrodiagnostic tests differ from each other. Laboratory tests usually cost less and are ordered more commonly, while radiological and electrodiagnostic tests are more expensive, complex and advanced (Janssens 2010; Shalev et al. 2009; Fuller 2005). The difference between the use of radiological test and electrodiagnostic test is that radiographic imaging is more commonly used than electrodiagnostic tests, and imagining is typically a prior examination, while electrodiagnostic test, especially ENMG, is often the extension to other examinations (Fuller 2005). Many studies have focused on reducing the number of laboratory tests, and characteristics used to determine unnecessary tests come from the field of laboratory (Fryer & Smellie 2013; Janssens 2010; Plebani et al. 2014). In this setting, inappropriate usage is explored from the too frequent retesting perspective (Janssens 2010; Fryer & Smellie 2013). Although too frequent control tests are also a problem in the field of electrodiagnostic testing, it is not the most common reason for insignificant tests. According to prior literature, a greater problem is that unnecessary ENMG test are those requested too early after the event of injury, while there are no findings possible to examine yet. In addition, ENMG test not extending clinical examinations or ENMG tests requested without other examinations are considered also insignificant tests. Furthermore, those tests that are not useful in the diagnostics since the symptoms of the patient are not related to neuromuscular diseases, are considered as inappropriate tests as well. (Di Fabio et al. 2013; Mondelli et al. 2014; Mondelli et al. 1998; Podnar 2005; Cocito et al. 2006.)

Nonetheless, there are conflicting opinions about whether a normal ENMG test result is insignificant (Shepherd 2010; Mondelli et al. 1998; Di Fabio et al. 2013; Cocito et al. 2006; Podnar 2005). Shepherd (2010) claims that a high rate of abnormal tests indicates that the test has been conducted on the right patient population, as many of them show findings of neuromuscular diseases. Furthermore, Di Fabio et al. (2013) concluded that normal ENMG tests were considered unnecessary when the patient's pre-test symptoms did not indicate any neuromuscular disease or peripheral nervous system dysfunction. Similarly, Podnar (2005) stated that a low rate of abnormal findings represents low quality of referrals, resulting from inadequate clinical examination before the referral.

On the other hand, Loop & Lusted (1978) presented that effectiveness of a diagnostic test should be measured as an optimal ratio of normal and abnormal results produced; a high ratio of normal results could indicate a high number of inappropriate tests. Yet, normal results can provide necessary and important information to the referring physician, such as excluding some diagnoses from a set of hypothetical diagnoses. While an abnormal result confirms a disease, a normal result excludes considered diagnosis and makes other possibilities worth considering (Cocito et al. 2006). Van Walraven & Naylor (1998) pointed out that if all normal results were considered as inappropriate tests, the appropriate use of diagnostic tests would require that a physician ordering a test should know the result in advance in order to avoid normal test results, making the whole test unnecessary. Due to these reasons, all normal diagnostic test results cannot be considered insignificant (van Walraven & Naylor 1998). Furthermore, van Walraven & Naylor (1998) claim that tests which have highly predictable results are unnecessary and insignificant.

There has also been criticism on the implicit definitions of insignificance of some prior studies on insignificant tests. Van Walraven & Naylor (1998) studied earlier publications on the subject and listed inappropriateness criteria used for laboratory tests they deemed implicit. They mainly listed criteria that were too unspecifically defined or were not evaluated from the disease perspective. On this list, van Walraven & Naylor (1998) included criteria such as "*tests which were categorized as inappropriate by clinical specialists*", "*tests that did not contribute to final diagnosis or treatment plan*" and "*tests which were normal*". In contrast, they present that proper, explicit inappropriateness criteria include those that consider the perspective to prior tests, were based on national guidelines or used inappropriateness criteria that was used in prior studies. (van Walraven & Naylor 1998). In other words, van Walraven & Naylor (1998) claim that diagnostic tests should not be deemed inappropriate in isolation from other tests, and that inappropriateness criteria should not depend on only one specialist's opinion about inappropriateness.

2.3.2 Characteristics of inappropriate test requests

Like a test can be insignificant, the diagnostic test request done by the referring physician can be inappropriate (Janssens 2010). Commonly inappropriate test requests fulfill some of the following criteria; requests that do not include all relevant information about the patient (Cocito et al. 2006; Di Fabio et al. 2013), only the symptoms of a patient are described but no referral diagnosis is included (Mondelli et al. 1998; Podnar 2005), the test request is done prematurely or too late concerning patient's symptoms, or the test request does not include the appropriate tests that are able to examine the patient's symptoms (Janssens 2010; Fryer & Smellie 2013). The greatest challenge with inappropriate test requests is that rejecting test referrals by just reviewing them is risky (Fryer & Smellie 2013). The referral may seem to be inappropriate (formulated incorrectly), however, the test may turn out to be necessary after having seen the patient. One may succeed to reduce the number of inappropriate referrals by educating the referring physicians about the appropriateness of test requests, and what to include in it (Fryer & Smellie 2013; Miyakis et al. 2006).

Especially in electroneuromyography testing, the test referral is a crucial factor affecting both success and significance of an ENMG test (Robinson 2000). Due to the limited possibility of having an dialogue with referring physician at the moment of performing ENMG test, the referring quarter should include all relevant information concerning the patient's disease and symptoms (Fuller 2005). This same pattern has been noticed in the field of radiology (Lysdahl & Hofmann 2009). An ENMG test should be seen as an extension after clinical examination, and it should not be performed without a preceding clinical examination– the ENMG test does not primarily provide the cause for the disease, but rather confirms precise symptoms (Fuller 2005). In ENMG test requests, question setting seems to the most important factor that affects the usefulness of ENMG test.

It is essential to include the referral diagnosis and referral query to ENMG test referral. An ENMG test is generally useful only when it focuses on examining a specific peripheral nervous dysfunction (Robinson 2000; Podnar 2005). So (2009) argued that an ENMG test request should be based on a suspected diagnosis; if a referring physician cannot formulate a referral diagnosis he or she should not request an ENMG test at all (Fuller 2005; So 2009). In practice, referring physicians tend to criticize ENMG test reports which do not give answer to their questions, although in the referral they might have asked too general questions, such as 'neurologic disease?' or 'weakness with pain?' (Smith 2003). Many prior studies have reported that referring diagnoses are often missing from the ENMG test requests, providing only symptom description rather than diagnosis suggestion (Di Fabio et al. 2013; Cocito et al. 2006). According to Cocito et al. (2006), making an ENMG test referral without specific suspected diagnosis is waste of time and resources, as neurophysiologists then have to do the time-consuming search for proper diagnosis hypothesis before the test. Furthermore, Fuller (2005) reminds that clinical neurophysiologists are a scarce and expensive resource, and should not be used ineffectively. In addition, if an ENMG test is requested with too general question, it may lead not only to wasting resources but also misdiagnosing the patient. An ENMG test with inadequate prior knowledge may either not focus on the actual problem of the patient, or emphasize findings that are not relevant to patient's condition (Mondelli et al. 2014; Fuller 2005).

Besides the referral formulation, test ordering behavior affects the quality and significance of ENMG test (Fuller 2005). Fuller (2005) presents two types of test ordering behavior: pragmatist and completist. A pragmatist critically considers all test to order and tries to formulate the most efficient diagnostic strategy as possible. In contrast, a completist is a practitioner who wants to exclude all possible diagnoses, and orders every possible test for the patient. (Fuller 2005), which causes 'over-investigation'

(Lysdahl & Hofmann 2009). Based on the aforementioned findings, general practitioners would seem to be more completist in their behavior, while specialists are more pragmatist (Fuller 2005). General practitioners tend to have lower correlation between the referral diagnosis and post-ENMG diagnosis than specialists (Di Fabio et al. 2013; Mondelli et al. 1998; Podnar 2005; Cocito et al. 2006; Mondelli et al. 2014). Prior studies of referral diagnosis and post-ENMG diagnosis correlation (Di Fabio et al. 2013; Mondelli et al. 1998; Podnar 2005; Cocito et al. 2006; Mondelli et al. 2014) are in line with this typology regarding the behaviors of general practitioners as completists and neurologists as pragmatists. The pragmatist might be seen as optimal behavior, however, such a way of ordering tests may lead to problems such as misdiagnosis if too few diagnostic tests are ordered. A completist may also face the problem of misdiagnosis, which is due to the excess number of tests ordered and the possible false positive test results (Fuller 2005). To conclude, a physician needs to be critical with test ordering. On the other hand, a physician should not be too critical and mean with ordering of tests, in order not to misdiagnose a patient due to limited diagnostic test information.

The most common referral diagnoses differed between general practitioners and specialist. In general practitioner's referrals it was usually carpal tunnel syndrome, and in specialists' referral polyneuropathy, radiculopathy or other pathology (Di Fabio et al. 2013; Mondelli et al. 1998). Because specialists seemed to have a better agreement of referral diagnosis and post-test diagnosis, a couple of studies suggested that if a GP suspects other diagnoses than carpal tunnel syndrome, the patient should be sent to a neurologist or a specialist for further examination (Mondelli et al. 2014; Mondelli et al. 1998). The agreement of referral diagnosis and post-test diagnosis was between 3 and 20 percentage points better in specialists' referrals than in GPs' referrals (Table 1) (Podnar 2005; Di Fabio et al. 2013; Mondelli et al. 2014; Mondelli et al. 1998; Cocito et al. 2006). According to Mondelli et al. (1998) the congruence of referral diagnoses and diagnoses formulated by neurophysiologist was lowest in general practitioners' referrals (47.7 % of all GPs' referral diagnoses) and highest in neurosurgeons' referrals (73%). Specialists and especially neurologist and neurosurgeons tend to formulate much more focused referrals to ENMG tests than general practitioners (Mondelli et al. 1998; Mondelli et al. 2014; Di Fabio et al. 2013; Podnar 2005). Naturally, a neurological specialization enables better identification of different nervous system dysfunctions, which is reflected in the referrals made by neurologist or neurosurgeons.

2.3.3 The outcomes of reducing insignificant tests

Reducing insignificant ENMG tests and referrals will improve the quality of ENMG testing and reduce waiting times (Podnar 2005; Cocito et al. 2006). As shown in Figure 7 the timing of test, test result, diagnosis and treatment all affect the patient outcome. Therefore shortening of the waiting list improves patient outcome due to better timing of test and faster treatment. When ENMG test requests are formulated in sufficient detail and appropriately, then neurophysiologist is able to perform a focused and correct ENMG test (Cocito et al. 2006). Then time and resources are saved, and more efficient reallocation of scarce resources is possible (Eichler et al. 2004). However, perhaps the greatest advantages of reducing inappropriate ENMG test arise when patients with actual need will have the test within shorter period of time (Podnar 2005), and patients who do not need the test will not experience discomfort due to unnecessary invasive EMG needle test (Kothari et al. 1995) or get false positive test results (Fuller 2005).

According to prior studies, the share of insignificant ENMG tests and inappropriate test requests is 17–28% of all ENMG tests performed (Table 1). The number of insignificant ENMG tests presents approximately fifth of all ENMG test performed annually (Di Fabio et al. 2013; Mondelli et al. 1998; Mondelli et al. 2014; Cocito et al. 2006). Even though the criteria to define inappropriate tests varied between the studies, the percentages were comparable, and this data might be used as reference across different CN laboratories. In addition, healthcare policies, regulations and funding affect physician's test ordering behavior, which affects the number of inappropriate tests requested (Eisenberg 2002). Therefore the number of inappropriate tests may vary between different countries. Also, the different research methods explain the variation in the percentage of inappropriate tests. Most of the prior studies used general data analysis, which refers to gathering information about e.g. patient's personal data, specialty of physician requesting the ENMG test, clinical information included in test referral, referral diagnosis and the finding from ENMG test (Di Fabio et al. 2013; Mondelli et al. 2014; Cocito et al. 2006; Shepherd 2010). Mondelli et. al. (1998) conducted a short survey to neurophysiologists concerning the referral diagnosis and its appropriateness.

However, insignificant ENMG tests are often linked to inappropriate ENMG test requests (Di Fabio et al. 2013; Cocito et al. 2006; Mondelli et al. 1998). Inappropriate test requests refer to ENMG test request without referral diagnosis or with too general query (Robinson 2000; Smith 2003), resulting in extensive ENMG examinations. As a result, reducing the number of inappropriate test requests will release much more resources than expected. Still, the number of requests without referral diagnosis or specific query varies between 21 % and 40 % of all referrals (Table 1) (Di Fabio et al. 2013; Mondelli et al. 1998; Mondelli et al. 2014; Cocito et al. 2006; Podnar 2005).

Table 1 Summary of prior studies of EMG and NCS tests' and test referrals' appropriateness

Study	Country	Type of test	Research method	no. of patients examined	Normal results (%) of all tests	% of tests ordered by GP	% of tests ordered by specialist	% of referrals with referral diagnosis	% of referral diagnoses confirmed	% of GP's referral diagnoses confirmed	%of specialist's referral diagnoses confirmed	Criteria for inappropriate tests	% of inappropriate tests
Di Fabio et al. (2013)	Italy	EMG	Data analysis [a]	1220	25 %	57 %	37 %	32 %	58 %	54 %	64 %	(1), (2), (3)	17 %
Mondelli et al. (1998)	Italy	EMG	Survey and data analysis [a]	2709	22 %	49.9 %	50.1 %	76.6 %	45.4 %	41.1 %	51.3 %	(1)	22.8 %
Mondelli et al. (2014)	Italy	EDX	Data analysis [a]	1586	37.5 %	66.2 %	33.8 %	65.1 %	71.9 %	70.9 %	73.4 %	(5), (4)	21% (could have avoided with careful clinical examination)
Cocito et al. (2005)	Italy	EDX	Data analysis [a]	3900	37.5 %	25 %	75 %	79 %	40.5 %	36.5 %	41.8 %	(8)	28 %
Podnar (2005)	Slovenia	EMG	Data analysis [a]	300	55 %	65 %	35.3 %	59.7 %	44 %	47.4 %	67 %	(2), (3), (5)	> 56 %
Shepherd (2010)	US	EDX	Data analysis [b]	100	12 %	100 %	0 %	100 %	-	-	-	(6), (7), (8)	(40 % did not confirm diagnosis, 70 % did not change clinical management
Karadag et al (2014)	Turkey	EDX	Data analysis [a]	2701	37.9 %	0.7 %	99.3 %	100 %	-	-	-	-	-

a Retrospective research method
b Before-after-study method

(1) did not confirm neurophysiologist's working diagnosis
(2) did not explain patient's symptoms
(3) test gave normal result and patient did not have symptoms of peripheral nervous dysfunction
(4) test referral was inaccurate without referral diagnosis
(5) test referral was not in relation to clinical findings
(6) test result was normal
(7) test did not change treatment plan
(8) test did not affect pre-diagnosis or treatment plan

To conclude, inappropriate testing has been presented as a major reason for increased diagnostic test utilization. Reducing these inappropriate tests would result in great savings. Minimizing the number of insignificant diagnostic tests can be reflected into other healthcare sectors as well, while ineffectiveness is most probably present also in other healthcare services from where insignificant procedures could be removed (van Walraven & Naylor 1998). In addition, inappropriate diagnostic test utilization generates unnecessary discomfort among patients (Thornbury et al. 1999; Kothari et al. 1995) and increases the number of false positives (Fuller 2005), which generate a burden to other healthcare sectors as well (van Walraven & Naylor 1998).

3 Research methodology

This chapter describes the empirical setting of this study, namely a case study research approach. Furthermore, the research design and the data collection and analysis methods are presented in more detail.

This study is a document-based, non-experimental research, in which data is gathered retrospectively from HUS patient information systems and from referring physicians and neurophysiologists via survey and from three selected HUS specialists via interviews. The gathered information and the results of this study did not affect the patients already examined. The present study uses a case study method (Creswell 2014). This study includes both qualitative and quantitative research methods. However, the emphasis is on the quantitative survey research method.

Statistical analysis methods are used to explore survey data, performance data and data gathered with closed-ended interview questions. The survey data gathered from neurophysiologists is analyzed with ordinal logistic regression model and chi-square tests. In addition, the differences between HUS CN laboratories, based on performance data, are statistically tested with Student's t-test (Taanila 2012). The present study uses both cross-sectional and longitudinal surveys (Creswell 2014) to gather data. Qualitative data analysis includes categorizing open-ended interview answers, which then are explicated and summarized. However, the semi-structured interviews of this study include mainly closed-ended questions that are gathered to support the results from surveys, additionally these closed-ended interview answers are analyzed with quantitative methods.

3.1 Research task

The aim of this study is to help Meilahti CN laboratory to understand the reasons behind insignificant test requests, and how and which ENMG tests to prioritize in order to reduce the number of insignificant tests. For the first, this study identifies the characteristics of insignificant ENMG test and inappropriate ENMG test referral. After the identification, possible demand management strategies searched from literature are proposed to reduce and control the number of these specific insignificant ENMG tests and test referrals. Lastly, the number of insignificant ENMG tests are calculated for Meilahti CN laboratory based on the answers gathered from two surveys. Albeit diagnostic tests would be good to analyze and evaluate side by side with other diagnostic tests, within the limits of this study, this was not feasible. It is both difficult, complicated and time-consuming, thus diagnostic tests are often evaluated in isolation (e.g. Mondelli et al. 1998; Lijmer et al. 2009; Cocito et al. 2006).

Based on these objectives of this study, following research questions are answered; "What strategies could be used to reduce the number of insignificant diagnostic tests in healthcare?", "What are the characteristics of insignificant ENMG test in CN laboratory in Meilahti hospital?" and "What number of

ENMG test requests and ENMG tests are insignificant in Meilahti CN laboratory and do different referring units differ in this?”. The main research question is explanatory while the other two supportive questions are more descriptive in nature.

To conclude, in order to answer these questions and to achieve the aim of this study, a literature review, two surveys, three interviews and performance data collection are conducted. The main data is gathered via surveys done to both referring physicians and neurophysiologists. The three interviews are made in order to gather information about two specific patient diagnosis groups and to get support for the results from statistical analysis of survey answers. Performance data is gathered from Meilahti CN laboratory’s patient information system in order to gain understanding of the current situation and the problem at hand. Furthermore, a summary of prior studies is constructed in order to review a solution to the increased ENMG test demand through various demand management strategies.

3.2 Research design and approach

3.2.1 Quantitative and qualitative research

According to Creswell (2014), a determination of proper research design for the study depends on the nature of research problem. Since this study aims to answer two descriptive research questions of “what are the characteristics” and “how many” (Pinsonneault & Kraemer 1993) and one explanatory question, the present research mixes qualitative and quantitative research methods. However, this research uses mainly statistical analysis methods and gathers numeric data (Creswell 2014), therefore the emphasis is on the quantitative research method, even though this study uses the most common qualitative research method – the case study approach (ibid.). With qualitative research method, the present study explores phenomenon from the individuals’ perspective within a natural setting (Malterud 2001).

3.2.2 A case study research approach

The case study method is suitable for this study, because this study focuses only on one case: Meilahti CN laboratory, and especially, to the significance of ENMG tests. Furthermore, this research explores the research problem through different lenses by gathering data from different sources, which is a strength of case study approach (Baxter & Jack 2008). In addition, according to Yin (2009), a case study method is the preferable method if the focus of the research is on a present phenomenon within a current and actual process and that the researcher is not able to control the studied event, if the research answers to exploratory or explanatory research questions and if the studied phenomenon is highly dependent on the context. Although not all of requirements of a case study research are met in the current study, it still fits the best with the aim of this study. Within this study the behavior of physicians cannot be controlled by the researcher and this study explores the actual behavior of physicians when requesting tests. Furthermore, the studied phenomenon, the significance of ENMG test and the reasons for ordering insignificant test, is highly dependent on the medical and clinical context. However, this study has a descriptive focus instead of exploratory as it tries to answer questions of "what are the characteristics" and "how many". Nonetheless, the findings of this study are generalized to answer to the first research question concerning whole healthcare sector, hence this research provides answer to explanatory question as well.

The reasons for case selection were the following; Meilahti CN laboratory was selected to be the focus of this study, because this laboratory has the highest volume of electro diagnostic tests within the HUS area that require actions to minimize these volumes in the future. In addition, ENMG test waiting times are long within the laboratory and needs to be shortened. Furthermore, ENMG test requires a lot of neurophysiologist's time to perform, therefore minimizing the number of inappropriate ENMG tests will save a lot of resources and labor time and provide cost savings to test referring quarter.

According to Yin (2009) a case study method uses multiple methods for data collection. The main data collection method used in the present study is survey, because it fits well with the characteristics of this study. According to Pinsonneault & Kraemer (1993) the survey research method is preferred method when the research fulfills the following requirements. First, the researched phenomenon is being studied in its natural setting. Second, the focus of the research has occurred in recent past or is currently happening. Third, the independent and dependent variables cannot be controlled or the research wants to keep the variables uncontrolled. Lastly, survey research approach is preferred if the research aims to answer research questions of "what" and "how many". (Pinsonneault & Kraemer 1993.) All of these requirements are fulfilled in this study.

One of the surveys used in this research is cross-sectional as it collects data in one point in time and the other survey is longitudinal survey, which collects data in two points in time (Pinsonneault & Kraemer 1993). Survey done to referring physicians is cross-sectional as it collects their opinion only once and in one point in time, namely while reviewing the final report of ENMG test results. However, the survey done to neurophysiologists is a longitudinal survey, as it is divided into two parts, first part is answered before and the second after preforming the ENMG test to patient.

3.3 Data collection

Three different data collection methods were used. First, the performance data from patient information systems of all CN laboratories within HUS area was collected and analyzed in order to formulate a whole picture of the problem at hand. Furthermore, the differences between these CN laboratories were explored in order to plan and focus future actions. One of the goals was to determine which ENMG tests are increased, how much compared to previous years and why. The analysis of performance data is used to answer the first research question concerning valid demand management strategies. In order to select and propose demand management strategies, one has to understand the main problems in CN laboratories' performances. Besides the demand management strategies, performance data was used to extend the information gathered with surveys, such as adding the information of the referring quarter to evaluated patient case.

Second, in order to answer the second research question about the characteristics of an insignificant ENMG test, a survey was conducted to the referring physicians and neurophysiologists concerning the significance of the requested ENMG test. A survey (Appendix B: Survey made for referring physicians) was carried to the referring primary and secondary care physicians in the area of Helsinki. Survey was conducted in Finnish (Appendix E: Kysely tilaajataholle). First, an electric survey was sent to referring physicians attached to the ENMG test report, but unfortunately it became clear that electric survey formula did not reach enough audience due to the lack of usage of electric and digital patient information reporting systems. Hence, top 10 ENMG test requester units were selected, which account approx. 50 % of the total ENMG test requested annually, and to these units a paper version of the same survey was sent. The survey was conducted during November to December 2014. A total 70 patient cases were evaluated by referring physicians from 10 different referring units.

In order to answer to the last research question about the number of insignificant ENMG tests, survey made to neurophysiologists was conducted. The survey (Appendix A: Survey made for neurophysiologists) was carried to neurophysiologists in Meilahti and Laakso CN laboratories. The survey to neurophysiologists was conducted in Finnish (Appendix D: Kysely KNF-lääkäreille) A printed survey was attached to ENMG test referral and neurophysiologist was compelled to answer the first half of the survey before examining or interviewing the patient before the actual test and the second part after performing the test. With this method it was possible to gather information about neurophysiologist's ability to evaluate the significance of ENMG test based only on test referral. The survey was conducted during November to December 2014. A total 282 patient cases were evaluated by 11 neurophysiologists.

The potential sampling bias in finding representative physicians was dealt with recruiting all neurophysiologists working in Meilahti CN unit to answer the survey within the selected time period. Selecting ten the most requesting referring units, gives good representative sample of referring physicians. Physicians were asked to answer the survey every time they requested or performed an ENMG test within the selected time period, an therefore this study has a good representative sample of patient cases.

These surveys gave the advantage to compare the opinions of referring physicians and neurophysiologists on similar patient cases, and therefore provided the answer to the second research question concerning the differences in opinions. Altogether 33 patient cases were evaluated by both sides. The test identification codes that made the opinion comparison possible was gathered from patient information systems.

The third data collection method was interviewing three clinical specialists, which helped to answer to the second research question. The first interviewee is a head of orthopedic and traumatology department in HUS Töölö hospital. Second interviewee is a chief physician of orthopedic and traumatology clinic in HUS Töölö hospital. Third interviewee is a head of orthopedic surgery department in Peijas hospital. These three interviewees have experience of requesting diagnostic tests and therefore they were able to provide viewpoints to which affects the significance of diagnostic test. The three specialists were specialized in spinal, musculoskeletal or foot surgery. For the interviews we selected approx. 20 ENMG patient cases to each interview. The patient cases were first selected by reviewing ENMG test reports and then ensured the representativeness of cases by reviewing the selected cases by neurophysiologist. In total 20 spinal cases were selected for spinal surgery specialists, 17 *meralgia paresthetica* cases were selected for musculoskeletal surgery specialists and 19 *morton neuroma* cases were selected for foot surgery specialist. Spinal cases were mostly requested by the interviewee himself. However, there were both cases of *meralgia paresthetica* and *morton neuroma*; some of the cases were requested by the interviewee himself and some by other physicians. Nonetheless, interviewees were encouraged not to focus on the fact that who had ordered the ENMG test, rather to whether the test provided any additional value to the referring physician while making decisions concerning the patient case. The meaning was to get the third general opinion about the ENMG tests' significance in addition to neurophysiologists' and referring physicians' opinions. In addition, the aim of these interviews was to gather information about the selected specific patient diagnosis groups, and whether performing a difficult ENMG test to these patients really provides new information or helps with diagnosing or treatment plan decision-making. The interviews performed had closed-ended questions (Appendix C: Interview questions to specialists), which were categorized and analyzed with quantitative content analysis methods. The interviews were made in Finnish (Appendix F: Haastattelukysymykset kliinikoille).

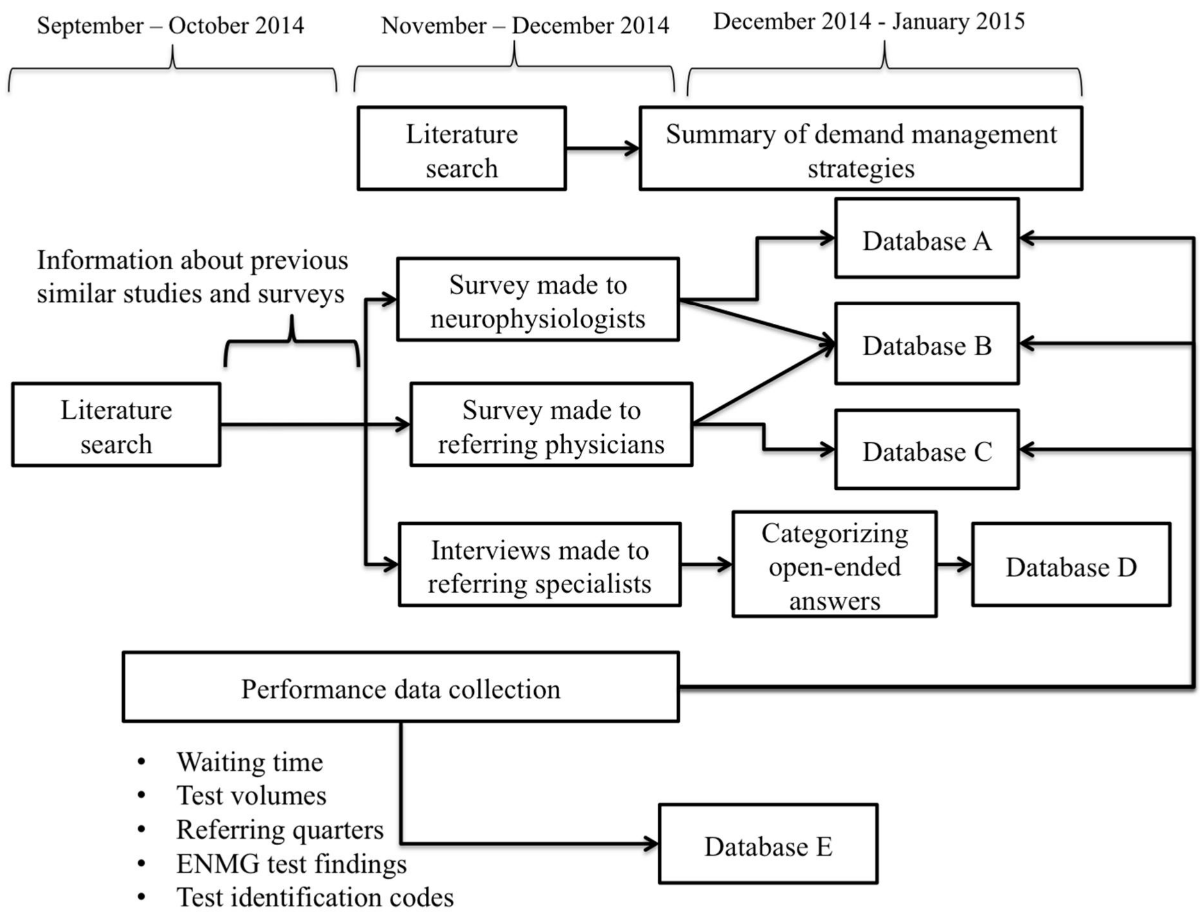

Figure 8 Data collection process of this study

3.3.1 Survey material

Both surveys were first formulated by the author of this study and then evaluated by two physicians, of whom the other one was the instructor of this study, in order to avoid questions that would have been either difficult or irrelevant to answer. In addition, both surveys included multiple-choice questions, therefore the answer alternatives were extremely carefully selected and then reviewed by experienced researcher in order to avoid answer alternatives that are either easy or tempting to select. The only open-ended questions were in survey done to neurophysiologists concerning the referral diagnosis and working diagnosis, which were limited to short answers i.e. a couple of words.

Survey material was collected from both neurophysiologists and referring physicians. Both surveys were conducted during November – December 2014. During that time all neurophysiologists working in Meilahti CN laboratory were required to fill up the survey every time they performed ENMG-1, ENMG-2 or ENMG-3 test. Altogether 11 neurophysiologists answered to the questions, four of whom were interns. However, not all referring quarters were involved in the survey done to referring physicians, as the first version of the survey was done electronically via HUS electronic reporting systems to all referring quarters, which ordered ENMG tests during that time, but failed to reach enough audience due to the minimal usage of such electronic reporting systems. Therefore a paper version of this similar survey was made to top 10 referring quarters, which accounted for 50 % of all ENMG tests requested annually. A paper survey was delivered to these units and at the same time a short introduction of this research was given to these quarters in order to ensure that referring physicians understood the aim of this study and to increase the answering rate.

The survey done to neurophysiologists was a paper attached to the ENMG test referral and it was divided into two sections, which were answered before the test and after the test. The aim of the survey was to gather information about neurophysiologists' opinion which ENMG test are insignificant and what factors affect the significance of a test. The survey included question of neurophysiologists' prior evaluation of the significance of ENMG tests, which was purely based on the evaluation of ENMG test referral, as neurophysiologists were asked not to interview the patient before answering the first part of the survey. The first part of the survey also included questions concerning the working diagnosis and referral diagnosis proposed, and whether neurophysiologist was certain about the working diagnosis. In addition, the survey had a question of whether the referral diagnosis was meant to either exclude or confirm the diagnosis proposed, and neurophysiologist was asked to evaluate the quality of ENMG test referral query. After performing ENMG test to patient, neurophysiologist answered to the second part of the survey. In the second part neurophysiologists provided the final evaluation of the significance of the preformed test and whether it provided new information. In addition, the neurophysiologist evaluated the extent of the test and the reasons for the significance of ENMG test. In order to analyze the survey answers, performance data was collected from patient information systems to fill up the missing information concerning the patient cases evaluated. The additional information included referring quarter, findings of ENMG test and extent of performed ENMG test. After this a comprehensive data table was built including both survey answers and performance data.

This study used the framework presented by Thornbury et al. (1999) to identify insignificant tests from referring physician perspective. According to Thornbury et al. (1999) a diagnostic test, that is not evaluated to be helpful with either diagnosing or treatment plan decision-making, is considered as less significant. Moreover, several other ENMG studies have supported Thornbury's framework (Fuller 2005; Di Fabio et al. 2013; Cocito et al. 2006; Podnar 2005; Rigler & Podnar 2007; Mondelli et al. 1998; Mondelli et al. 2014).

The survey done to referring physicians was meant to be answered when a physician received the report of ENMG test results and informed patient concerning the results. During or after the appointment where the results of ENMG tests were discussed with the patient, the referring physician answered to the survey questions. Survey included questions concerning the significance of ENMG test to patient diagnosis and treatment plan decision-making. In addition, in a survey referring physician was asked to evaluate the referral query they had included to referral and possible reasons for the significance of ENMG test. The final questions concerned referring physician's specialty and working experience. A database was build from the survey answers. Referring physicians were asked to write the ENMG test code, which identified the test from patient information system. This made the comparison of the opinions of referring physicians and neurophysiologists possible concerning the same patient cases.

3.3.2 Interview material

This study included three interviews made to three specialists in HUS with long working experience in the field of medicine. The meaning of these interviews was to gather information concerning two different patient diagnosis groups, to whom ENMG test is difficult to conduct and the outcome of such test is debatable. These patient diagnosis groups were *meralgia parestetica* and *morton metatarsalgia*. In addition, a spine surgeon evaluated patient cases concerning back problems in order to gather opinion in general about the significance of ENMG tests.

Interviews were held between November and December 2014 at the same times as surveys with the other instructor of this study, who has a medical education, which ensured that interview answers were understood by the author of this study. All of these interviews included approx. 20 patient cases, that had already been examined with ENMG test and the final reports had been written and diagnoses and treatments had already been decided. Interviewees were asked three questions concerning the patient cases and the significance of ENMG test in the specific case. In addition, these interviewees were asked a couple of general questions about the significance and meaning of ENMG test in general in diagnosing and treatment plan decision-making. The questions concerning specifically patient cases were closed-ended questions with multiple-choice questions, while the more general questions were open-ended. To each interview, information about the selected 20 patient cases were printed including the ENMG test referral, final report of ENMG test result and the final decision made by the referring physician based on the ENMG test result. This information was reviewed to interviewee and based on that information he evaluated whether ENMG test was significant of not, and what was the reason for the significance. Interviews were approximately one hour long and the answers were both written and recorded for the analysis.

3.3.3 Performance data

The secondary data of CN laboratory's performance was gathered from HUS medical imaging center's patient information system QPati. The information was collected from years 2011-2014, in order to see the trend of ENMG test volumes during the past years and also to analyze the differences between the CN laboratories within HUS area. Information gathered included identification code of test, typology of test, date of test referral received, date of test performed, referring quarter, CN laboratory where test was performed, test result and name of neurophysiologist. The numeric data was collected from the information system with different filters and exported to Excel file.

3.4 Data analysis

The data analysis of this research includes analysis of both qualitative data and quantitative data. The qualitative data was analyzed with qualitative content analysis method and quantitative data was analyzed with three different statistical methods: Ordinal logistic regression model, chi-square test and Student's t-test.

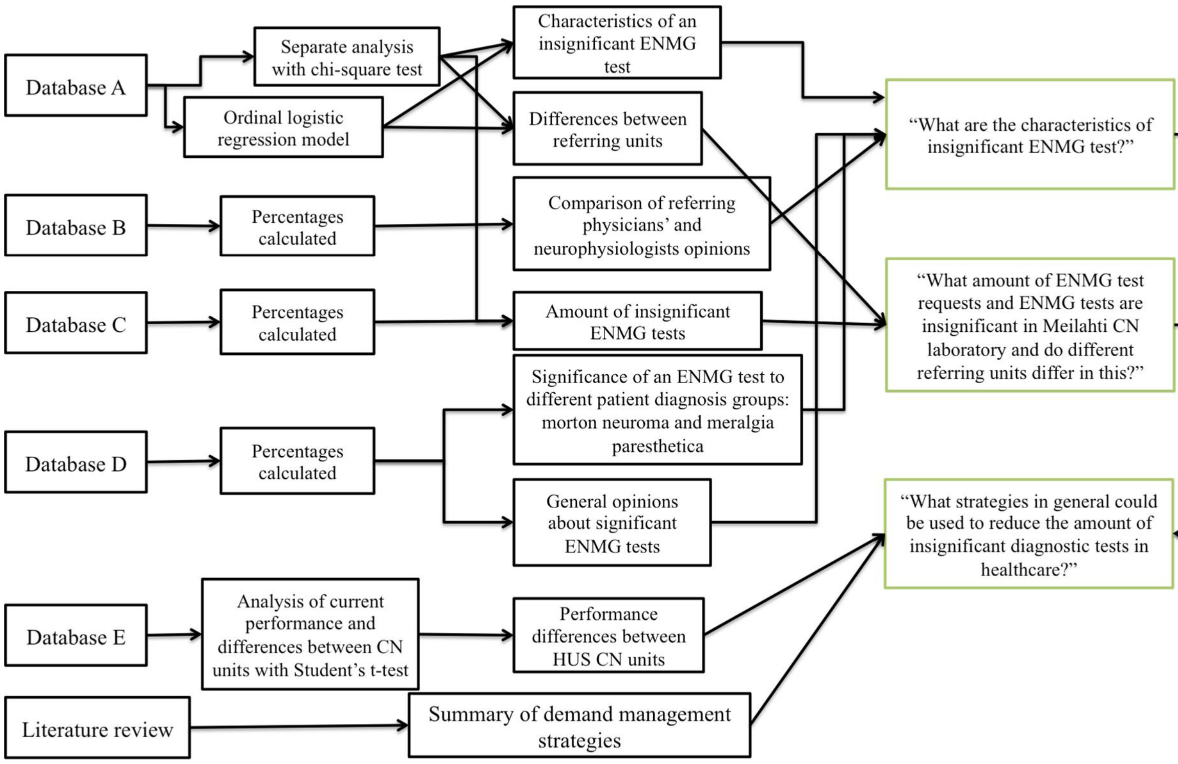

Figure 9 Data analysis process of this study and linkage to research questions

3.4.1 Qualitative data analysis

Qualitative data analysis is used for exploring the interview material of this study, as according to Creswell (2014) open-ended questions in interviews needs to be considered as qualitative data. This method includes gathering the qualitative data through interviews with notes and recordings. The analysis itself includes transcribing the audio recording, finding repetitive answers to the open-ended interview questions and then coding the answers from interview into numeric terms and then analyzing it with quantitative data methods. This analysis method is called content analysis approach (Krippendorff 2003). Content analysis approach is mainly used to making interferences from newspapers or other texts to context of their use (Krippendorff 2003) and to actually count the number of i.e. repetitive words within the text to analyze it with more quantitative method (Lacey & Luff 2007).

The qualitative data of interviews within this research was meant to support the findings of survey material analysis and provide additional information concerning the selected patient diagnosis groups. The interview was semi-structured having mainly closed-ended multiple-choice questions, which were coded numerically. However, the interview included a couple of open-ended questions, which needed transcribing and additional interpretation and coding of repetitive answers. The first open-ended question was “For what reason ENMG test is evaluated as insignificant or very insignificant?”. This question was presented to all three interviewees and the answers to this question were coded, categorized and used as support for findings of the survey data analysis. The other open-ended questions were presented to only one interviewee who was expert in the field. Due to the fact that there was only one answer to the questions concerning the patient diagnosis groups of *meralgia paresthetica* and one answer concerning *morton neuroma*, the answers were not coded numerically, but presented as such to support and reason the final result of statistically analyzed closed-ended multiple choice questions concerning the significance of ENMG test when performed on a patient with *morton neuroma* or *meralgia paresthetica*.

3.4.2 Statistical methods

The main data was gathered through surveys from neurophysiologists and referring physicians. The survey included mainly multiple-choice questions, which were coded into data table and the answers to open-ended questions were first categorized and then coded to the same data table. The data table was constructed into Excel work sheet having the questions in columns and each individual patient case in each row. The patient cases evaluated by neurophysiologists were categorized according to referring quarters. Referring units were divided into the following five groups; “hand surgery and surgery and orthopedic outpatient clinics”, “physiatry outpatient clinic and rehabilitation center”, “neurology and neurosurgery outpatient clinic”, ”other outpatient clinics” and “primary care”. “Other outpatient clinics” included e.g. rheumatology outpatient clinic, emergency rooms and wards. Finally, from the patient information system the test findings of ENMG were gathered and added into the data sheet with neurophysiologists answers in order to compare the proposed referral diagnosis, working diagnosis with the final finding of ENMG test.

To identify the main factors affecting the significance of an ENMG tests, an ordinal logistic regression model was built within a standard computer program (SPSS; IBM, Armonk, NY, United States) from the survey data gathered from neurophysiologists. Logistic regression model is used to explore and analyze the interaction and effects of selected variables (Sykes 2007). A regression model tries to model the effect of independent variable on dependent variable. An ordinal logistic regression model is a type of logistic regression model, in which the dependent variable is ordinal (Mccullagh 1980). Ordinal variable means that the variable it is not continuous scalable metric only it has stepwise values of i.e. 1 to 5, which represent “strongly disagrees” (1) to “strongly agrees” (5). Ordinal logistic regression models are usually used for modeling the findings of surveys especially variables, which are representing opinions. (Bender & Grouven 1997.) In addition to ordinal logistic regression model, Pearson and Spearman’s correlations were calculated for the variables with a standard computer program (SPSS; IBM, Armonk, NY, United States). Spearman’s correlation was used, as the majority of variables were non-parametric ordinal or nominal variables (Taanila 2012). With the Spearman's correlation it was explored whether there is a strong correlation between selected independent variables (correlation >0.5 or <-0.5), because then these correlating independent variables could not have been selected to the same model to explain the dependent variable due to multicollinearity.

Figure 10 presents the linkages between variables of ordinal logistic regression model, where the unit of analysis was a patient to whom an ENMG test was performed. The dependent variable was set to be the evaluation of final significance of ENMG test. Independent variables were selected from the different questions presented in neurophysiologists' survey. These independent variables included the quality of referral query, prior evaluation of significance, normal or abnormal ENMG test result, surprising or unsurprising ENMG test result, intern or specialist who performed the test, specialty of referring quarter, evaluation of the extent of a test, neurophysiologist's certainty of working diagnosis, extent of performed test and the agreement between referral diagnosis and final ENMG finding.

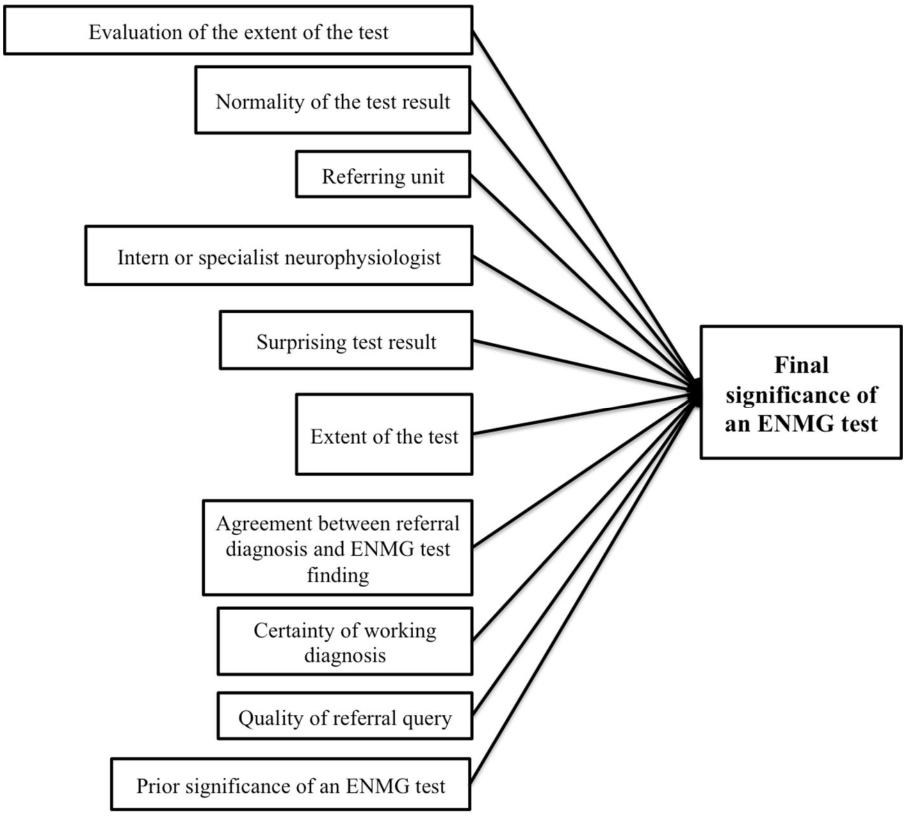

Figure 10 Ordinal logistic regression model variables

In addition to ordinal logistic regression model a separate statistical analysis of neurophysiologists' answers was calculated using chi-square test. The goal was to identify in more detail, which variables affect the significance of an ENMG test and especially how the referring quarter affects the significance. With chi-square test the differences of groups can be measured with categorical variables (Taanila

2012). Chi-square test is also used to measure the dependency between two variables (ibid.). Furthermore, the reliability of the data was tested with chi-square tests, in which were distribution of significance evaluations was compared between neurophysiologists.

The answers to survey done to referring quarter were not statistically analyzed due to the small number of patient cases evaluated. The total number of 70 patient cases was not large enough sample size to measure statistically. Furthermore, the comparison of opinions of neurophysiologists and referring physicians resulted in 33 patient cases, therefore no statistical analysis was done either as the sample size was too small. These answers were evaluated only by calculating the percentages of answers.

Lastly, performance data of CN laboratories was gathered in order to explore changes in ENMG test volumes and to have an understanding of performance of CN laboratories within HUS area, such as waiting time, test volumes, percentages of referring quarters. A Student's t-test was used to identify whether CN units statistically differentiated from each other.

4 Results

This chapter presents the results of qualitative and quantitative data analysis. Firstly, this chapter will present the current state of Meilahti CN laboratory. Secondly, the results of the input from neurophysiologists and referring physicians are presented in order to determine which kind of ENMG test are being considered as insignificant. Finally, this chapter presents the input from three interviews and the findings concerning the controversial diagnosis groups; *morton neuroma* and *meralgia paresthetica*. The sub-chapter 4.1 concentrates on describing the phenomena of increasing ENMG test volumes, while the last two sub-chapters will present the results of analysis of where to focus efforts of demand management.

This chapter is divided in terms of which are the main reasons for ordering an insignificant ENMG tests. Research material was gathered from patient information systems, from neurophysiologists and referring physicians with two different surveys presented in chapter 3.3.1 and from three interviews performed to specialists presented in chapter 3.3.2.

4.1 Current situation in Meilahti CN laboratory

During the years 2011-2014 the volume of ENMG test has increased by 26 % (Figure 11) in Meilahti CN laboratory. ENMG test volumes have increased annually by approx. 8.0 %. At the same time waiting time for ENMG test has become approx. 12 days longer 2014 vs. 2013 (Figure 12) in Meilahti. Furthermore, Meilahti CN laboratory has not achieved its strategic goal concerning the optimal length of waiting time except for two months in year 2014. The objective for waiting time of ENMG examination is to have at least 65 % of all ENMG tests to be done within 50 days or less, which days are calculated from when test referral is received to when test result report is send to referring physician. Also the other units have noticed the same problem of long queues.

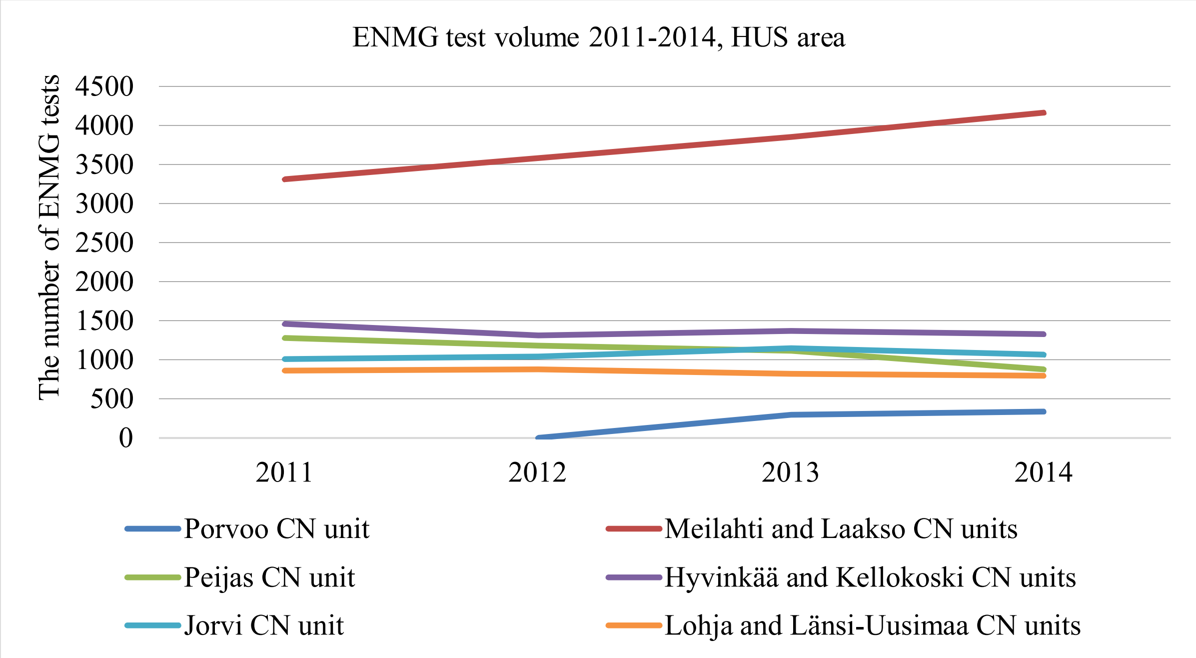

Figure 11 the development of ENMG test volume during years 2011-2014 in HUS area

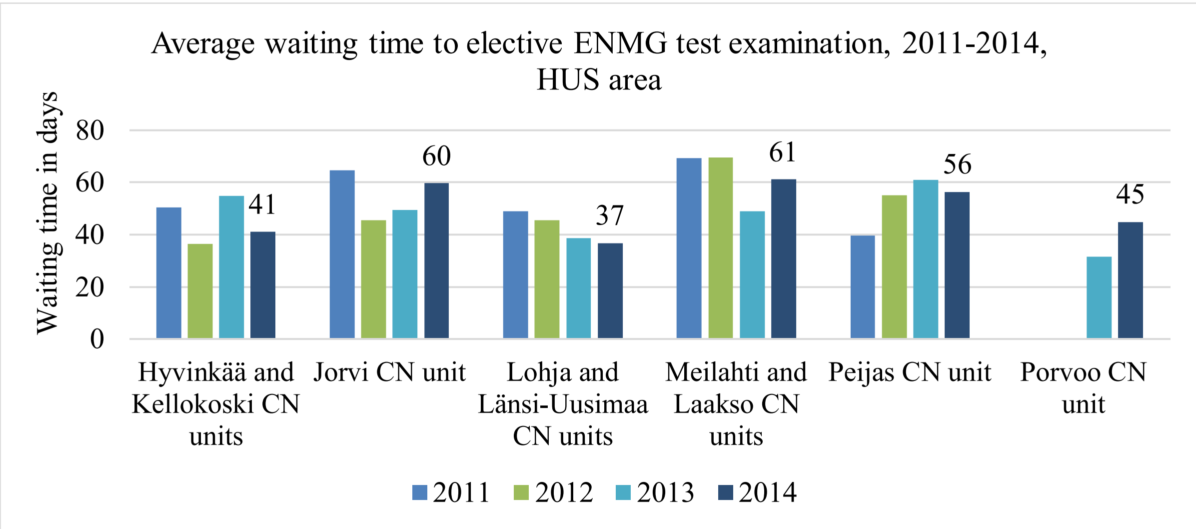

Figure 12 the development of non-emergency waiting time to ENMG test during years 2011-2014 in HUS area

CN laboratories within HUS area all differ from each other concerning ENMG test volumes, waiting time and specialty of test requestors. The greatest difference is in the profile of referring quarters (Figure 13). The percentage of ENMG tests ordered by primary care and physiatry outpatient clinic varies between CN laboratories. Due to these differences, CN laboratories may need to acquire different demand management strategies. Furthermore, the length of waiting time varies between Meilahti CN laboratory and all HUS CN units (t-test p-value 0.04) and especially between Meilahti CN unit and Hyvinkää CN laboratory and Lohja CN laboratory (t-test p-value 0.02). Waiting time is the longest in Meilahti, Jorvi and Peijas CN laboratories and shortest in Lohja and Hyvinkää (Figure 12).

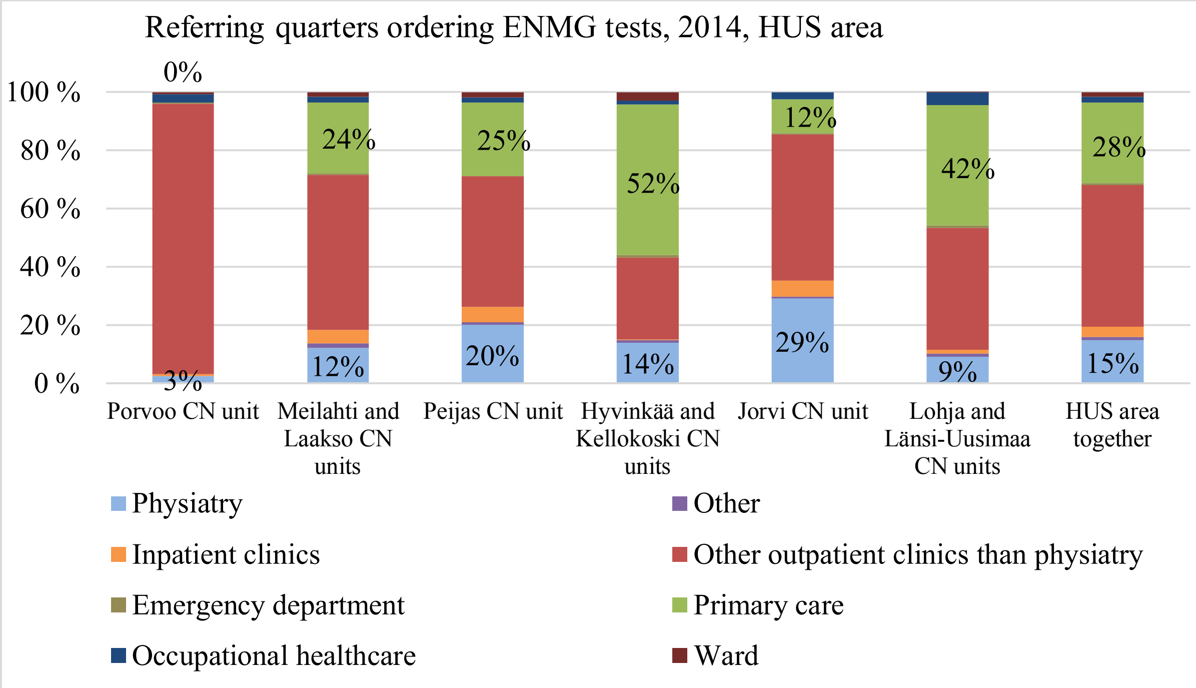

Figure 13 Referring quarters requesting ENMG tests in 2014 in HUS area

During the data gathering process of this study a great challenge in the communication with referring quarters was noticed. Normally, the only communication occurs when neurophysiologist reads an ENMG test referral, the referring physician reads ENMG test result report or when neurophysiologists call to referring physician directly about the test result in a case of emergency. In Meilahti CN laboratory they also arrange meetings with referring quarters. Specialists from different fields give lectures to each other about requesting and utilizing electrodiagnostic tests, in order to understand each other's view-points. However, this might not be enough: when a referring physician makes an inappropriate test re-ferral, he or she might get feedback 3-4 months later and not personally but in a summarized form, if at all – not every CN laboratory arrange these educational meetings with referring quarters.

In addition, there were no digital communication channels, through which these referring physicians would have found the survey used in this study or through which such a survey could have been sent to referring quarters. Therefore a paper version of the same survey was made and personally taken to these different quarters. This drawback during this research process represents well the current situation of challenging communication and the lack of effective communication channels in the healthcare sector, which can be considered as one of the findings of this study.

Hence, better and more agile communication channels are needed in order to improve the quality of referrals and selection of patients who benefit from ENMG testing. A better communication would help both service provider and orderer to manage their costs. Especially, the communication between CN laboratory and primary care is minimal. One would assume that sending emails is effective way to com-municate, but due to the constantly changing personnel in primary care, sending emails is not efficient. Due to this more effective real-time communication should be encouraged. Using phones is difficult, because physicians are extremely difficult to reach with phones as they are busy with their daily work, and they do not work at one place only.

4.1.1 Increasing volume of ENMG tests and the main referring quarters

The percentage of ENMG tests of all electrodiagnostic tests performed during year 2014 was ~45 % in Meilahti CN laboratory. The percentage has increased during last years being 38 % of all electrodiagnostic tests in 2011 (Figure 15). The overall volume of electrodiagnostic tests in Meilahti CN laboratory has increased by ~6.4 % between years 2011-2014 (Figure 14). Total increase of ENMG tests' volumes within whole HUS area was 8.3 % between years 2011-2014.

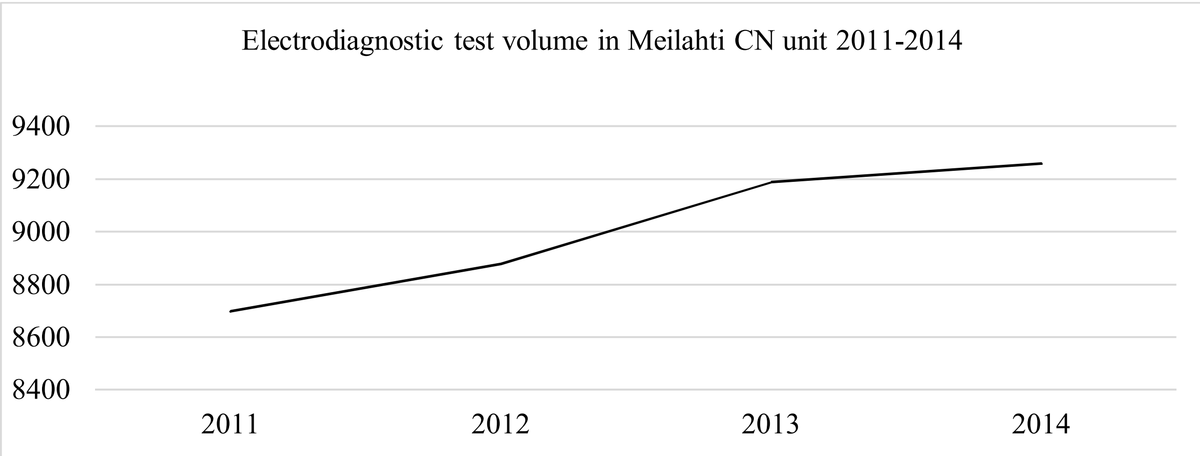

Figure 14 the development of electrodiagnostic test volume in Meilahti CN laboratory during years 2011-2014

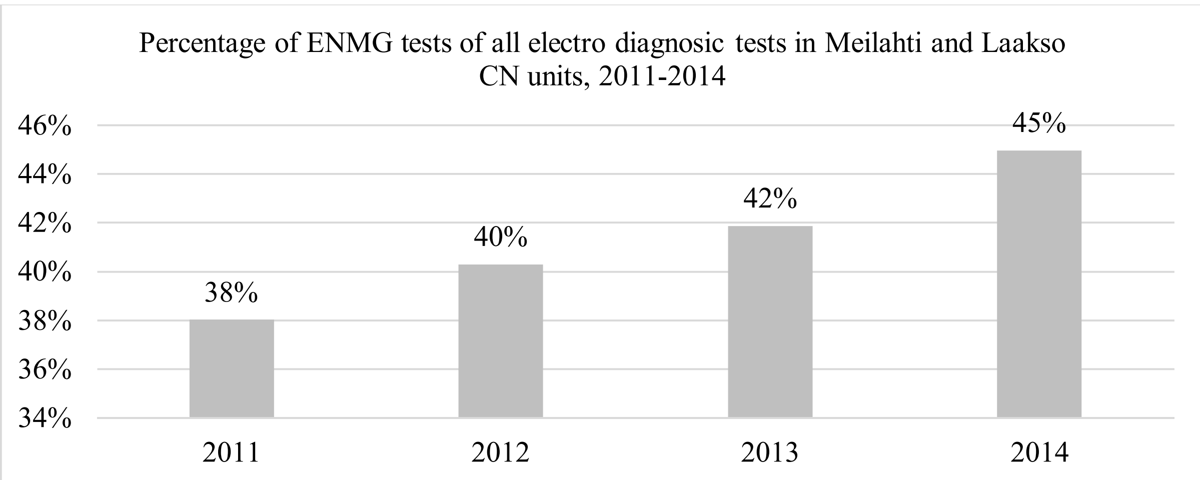

Figure 15 the development of the percentage of ENMG tests of all electrodiagnostic tests in Meilahti CN laboratory during years 2011-2014

Furthermore, the volume of ENMG-3 test in Meilahti CN laboratory has increased by ~160 % between years 2011-2014 (Figure 16). At the same time the volume of ENMG-1 test has decreased, while volume of ENMG-2 has slightly increased (Figure 16). In Meilahti, a majority of ENMG-3 tests are ordered by outpatient clinics. About 15 % of ENMG-3 tests are ordered by primary care physicians and ~11 % by physiatrists. In contrast, in Hyvinkää CN laboratory ~46 % and in Lohja CN laboratory 40 % of ENMG-3 tests are ordered by primary care physicians. (Figure 17.) In Jorvi CN laboratory and Peijas CN laboratory physiatrists order a quarter of ENMG-3 tests. The greatest difference between CN laboratories is the percentage of ENMG tests ordered by primary care physicians and physiatrists. In Meilahti they account 26 % of all ENMG-3 tests ordered annually, while in Hyvinkää and Lohja CN laboratories these referring quarters account for approx. 47 – 63 % of all ENMG-3 tests ordered (Figure 18).

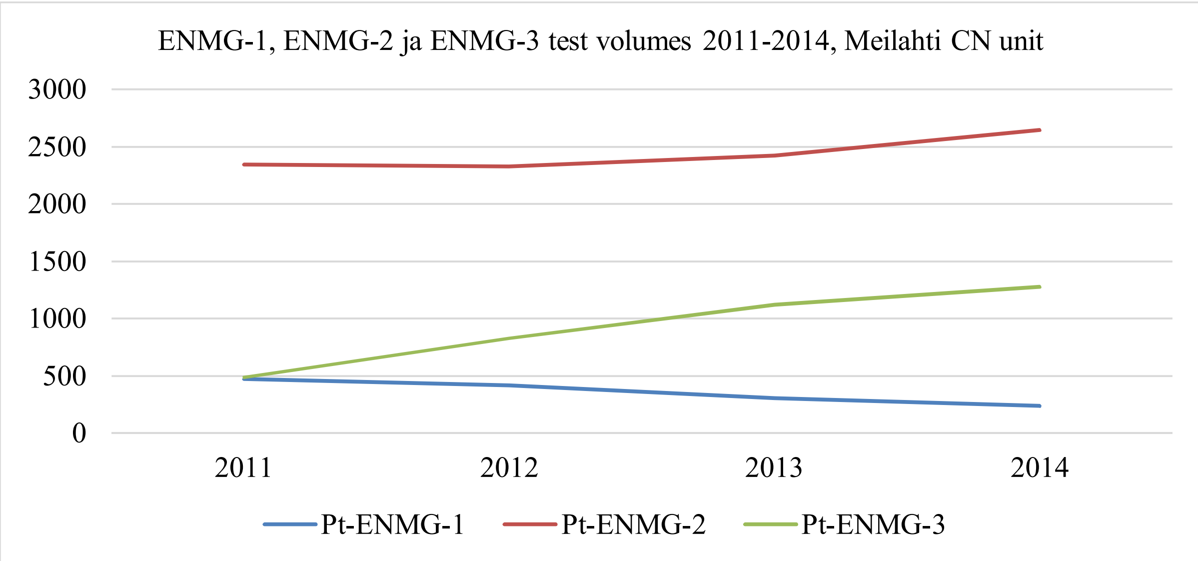

Figure 16 the development of ENMG-1, ENMG-2 and ENMG-3 test volumes in Meilahti CN laboratories during years 2011-2014

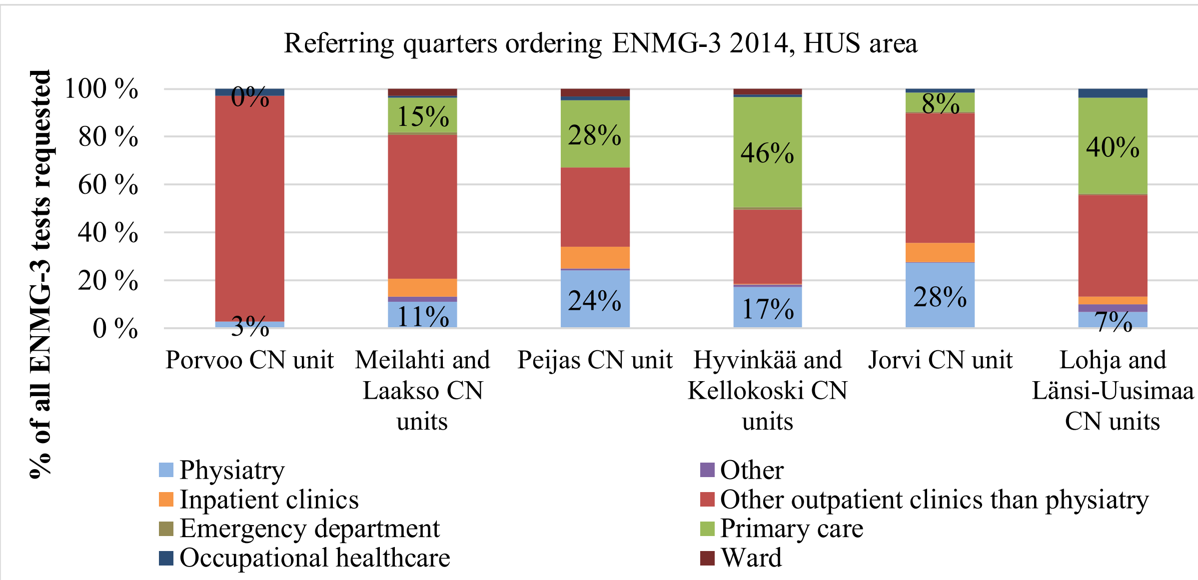

Figure 17 The percentage of ENMG-3 tests ordered by different referring quarters in HUS area in 2014

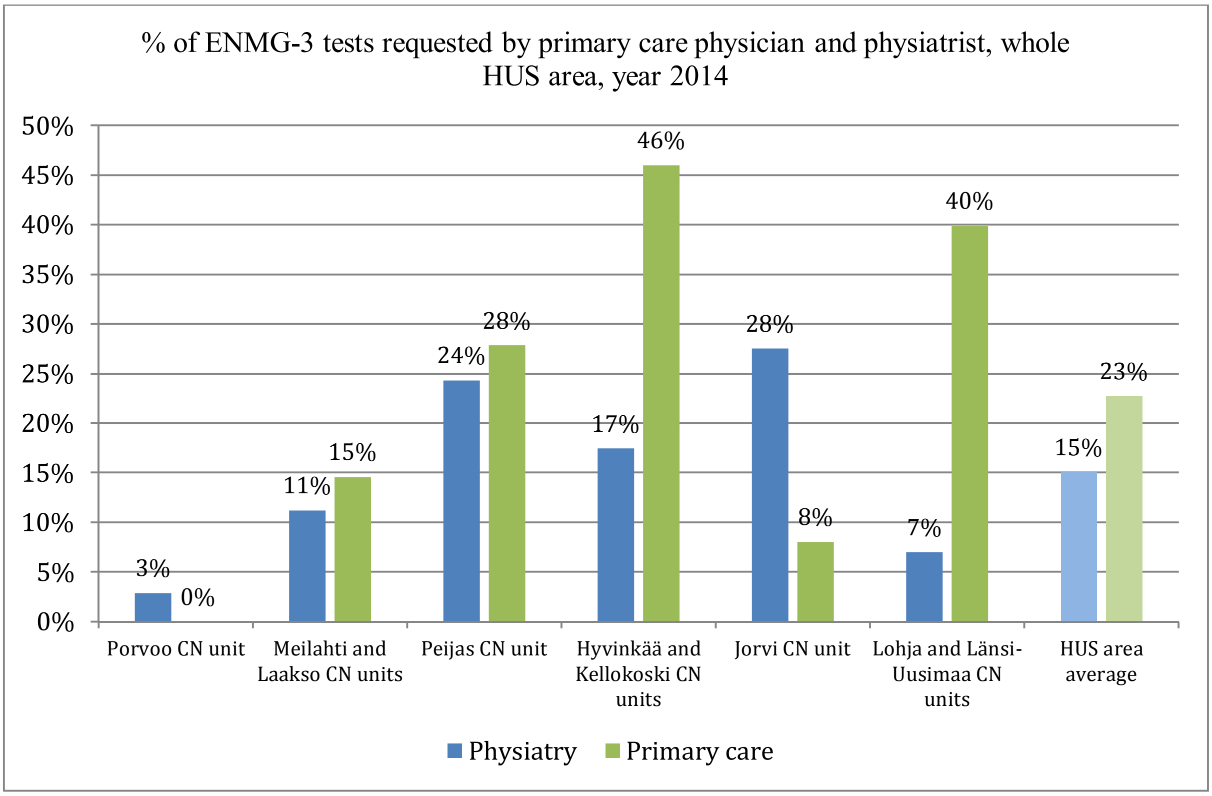

Figure 18 The percentage of ENMG-3 tests ordered by primary care physicians and physiatrist in HUS area, in 2014

4.1.2 Current actions used for shortening the waiting times

Meilahti CN laboratory has implemented actions to shorten waiting time for ENMG test and to reduce the costs. HUS *Medical Imaging Center* has aimed at decreasing the number of outsourced physicians, thus Meilahti CN laboratory reduced it by 75 % between years 2011-2014 (Figure 19). In 2011 the cost of outsourced neurophysiologists in Meilahti CN laboratory was 114 150 € and in 2014 the cost was only 28 250 € (Figure 19). Meilahti CN laboratory replaced outsourced physicians with more permanent personnel – neurophysiology interns. Because reducing outsourced neurophysiologists did not shorten the waiting time for ENMG tests, Meilahti used other methods. These methods included increasing the number of ENMG tests done outside office hours by permanent neurophysiologists, increasing the resource utilization by 4.5 % and implementing Lean methods (Womack et al. 1992). ENMG tests performed after office hours increased by 655 % between years 2011-2014 (Figure 20). However, the number of ENMG tests performed during office hours has not increased (Figure 21). This is a result of adding time-consuming new electrodiagnostic tests to the test repertoire, which new tests have taken time from ENMG tests. In addition to this, Meilahti CN laboratory has increased the capacity utilization rate by 4.5 % between years 2013-2014 measured in examination hours performed per employee's yearly working days.

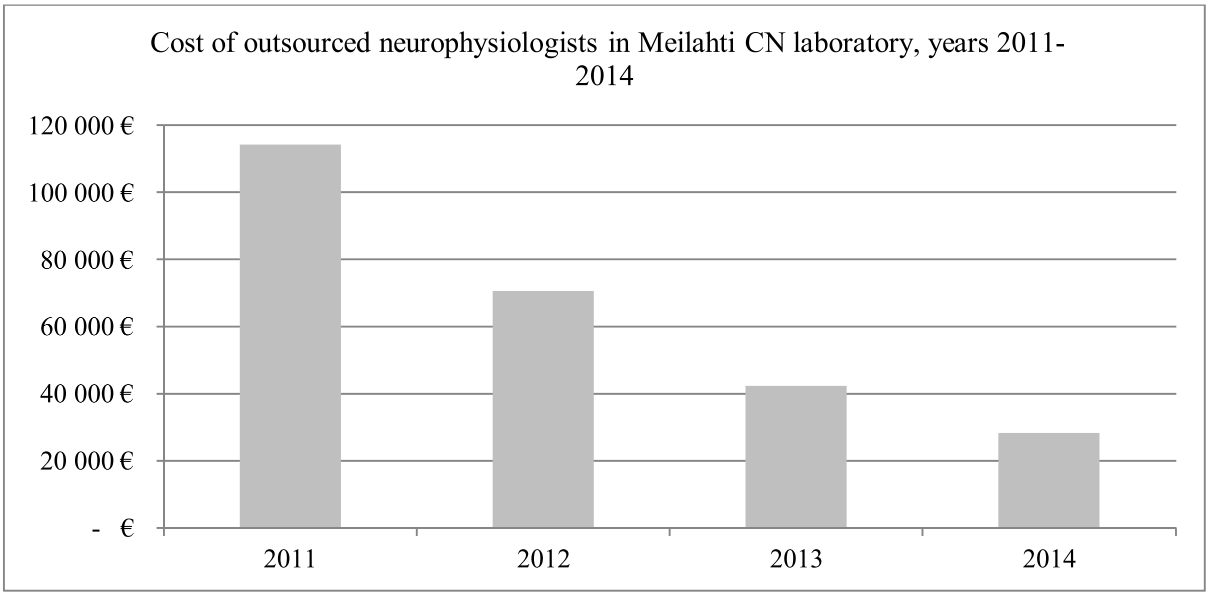

Figure 19 the development of cost of outsourced neurophysiologists during years 2011-2014 in Meilahti CN laboratory

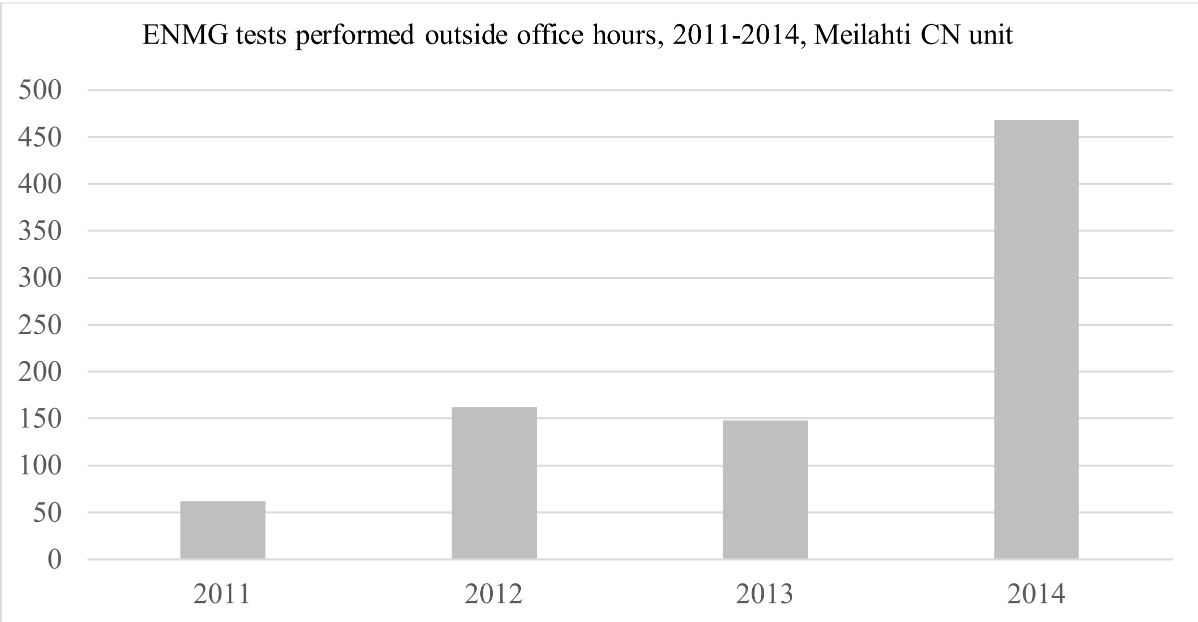

Figure 20 the development of ENMG test performed outside office hours in Meilahti CN laboratory during years 2011-2014

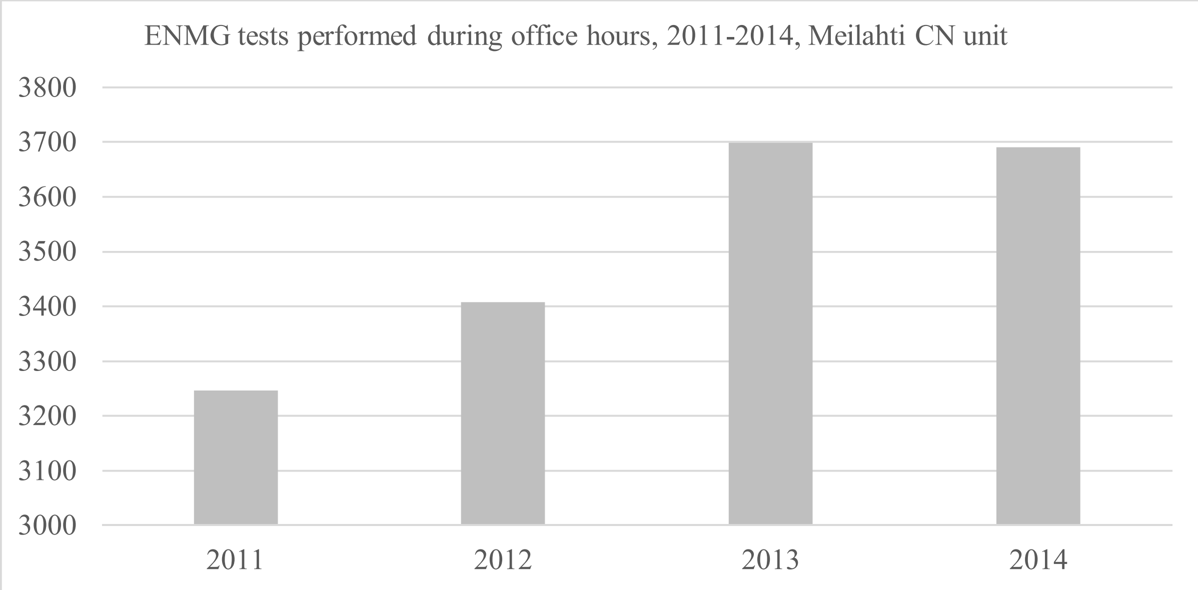

Figure 21 the development of ENMG tests performed during office hours in Meilahti CN laboratory during years 2011-2014

4.2 The main findings from surveys

The ENMG tests evaluated by neurophysiologists were requested by primary care in 28 % (n=78) of cases, and the rest 72 % (n=204) by secondary care physicians. The ENMG tests requested by secondary care physicians were ordered by hand surgeons and orthopedists (33 %, n=67), neurologists and neurosurgeons (28 %, n=58), physiatrists (18%, n=36) and other specialties e.g. rheumatologists, cardiologists, from emergency rooms and wards (22 %, n=45).

An ordinal logistic regression model was built in order to determine the main factors influencing significance of ENMG test. According to Pearson correlations (Appendix H: Pearson correlation between ordinal logistic regression model variables) none of the independent variables correlate strongly with each other, therefore these selected variables can be used in this regression model without any singularities and multicollinearity. The strongest correlation between independent variable and dependent variable is between prior significance and final significance (Pearson correlation 0.535, sig. 0.00). The second strongest is how surprising the ENMG finding is (Pearson correlation 0.380, sig. 0.00) and the third is whether ENMG was performed by intern or specialist (Pearson correlation 0.259 sig. 0.00). The definitions of variables used in the ordinal logistic regression model are shown in Table 3. The output from the best describing ordinal logistic regression model is presented in Table 2.

Table 2 Associations between independent variables and test significance evaluated by neurophysiologists. Ordinal logistic regression model – parameter estimates for independent variables, SPSS (N=263)

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Thresh-old	[Final significance = 1]	-7.489	.909	67.926	1	.000	-9.270	-5.708
	[Final significance = 2]	-7.187	.899	63.851	1	.000	-8.950	-5.425
	[Final significance = 3]	-2.899	.782	13.736	1	.000	-4.432	-1.366
	[Final significance = 4]	-2.473	.775	10.173	1	.001	-3.993	-.953
Location	[Prior significance=1]	-3.934	.576	46.589	1	.000	-5.064	-2.804
	[Prior significance=2]	-1.977	.450	19.331	1	.000	-2.859	-1.096
	[Prior significance=3]	0[a]	.	.	0	.	.	.
	[Referral query=1]	-.817	.612	1.782	1	.182	-2.016	.383
	[Referral query=2]	-.445	.351	1.606	1	.205	-1.133	.243
	[Referral query=3]	0[a]	.	.	0	.	.	.
	[Certainty of working diagnosis=1]	1.852	.962	3.708	1	.054	-.033	3.737
	[Certainty of working diagnosis=2]	.811	.530	2.341	1	.126	-.228	1.851
	[Certainty of working diagnosis=3]	.906	.509	3.160	1	.075	-.093	1.904
	[Certainty of working diagnosis=4]	0[a]	.	.	0	.	.	.
	[Evaluation of extent=1]	.126	.534	.056	1	.813	-.920	1.172
	[Evaluation of extent=2]	0[a]	.	.	0	.	.	.
	[Surprising findings=1]	-1.516	.416	13.258	1	.000	-2.332	-.700
	[Surprising findings=2]	-1.715	.602	8.104	1	.004	-2.895	-.534
	[Surprising findings=3]	0[a]	.	.	0	.	.	.
	[Extent of test=1]	-.179	.863	.043	1	.836	-1.869	1.512
	[Extent of test=2]	-.113	.345	.107	1	.744	-.789	.563
	[Extent of test=3]	0[a]	.	.	0	.	.	.
	[Specialist intern=1]	-1.197	.306	15.271	1	.000	-1.797	-.596
	[Specialist intern=2]	0[a]	.	.	0	.	.	.
	[Abnormal normal=1]	-.841	.357	5.529	1	.019	-1.541	-.140
	[Abnormal normal=2]	0[a]	.	.	0	.	.	.
	[Final vs referral diagnosis=1]	.239	.369	.421	1	.517	-.483	.962
	[Final vs referral diagnosis=2]	0[a]	.	.	0	.	.	.
	[Referring quarter=1]	-.138	.431	.103	1	.748	-.982	.706
	[Referring quarter=2]	.004	.478	.000	1	.992	-.931	.940
	[Referring quarter=3]	-1.374	.529	6.750	1	.009	-2.410	-.337
	[Referring quarter=4]	-.329	.441	.555	1	.456	-1.193	.536
	[Referring quarter=5]	0[a]	.	.	0	.	.	.

Link function: Logit.

a. Parameter is set to zero because it is redundant.

Table 3 Ordinal logistic regression model variable definitions

	The number and definition of variable				
Threshold and locations	**1**	**2**	**3**	**4**	**5**
1 Final significance	Insignificant	Slightly significant	Somewhat significant	Significant (needed for documentation)	Very significant
2 Prior significance	Slightly significant	Somewhat significant	Very significant		
3 Referral query	Referral query missing or unclear and poorly defined	Somewhat defined, but unclear or well-defined, but irrelevant	Well-defined, relevant and clear		
4 Certainty of working diagnosis	No working diagnosis	Not certain	Almost certain	Very certain	
5 Evaluation of extent	Too narrow or too broad	Appropriate			
6 Surprising finding	Unsurprising finding	Something surprising, but it did not affect patient diagnosis nor treatment plan decision-making	Surprising finding		
7 Extent of test	ENMG-1	ENMG-2	ENMG-3		
8 Specialist intern	Intern	Specialist			
9 Abnormal normal	Normal	Abnormal			
10 Final vs referral diagnosis	Different	Same			
11 Referring quarter	Primary care	Other outpatient clinics	Physiatry outpatient clinic and rehabilitation center	Hand surgery and general surgery and orthopedic outpatient clinics	Neurology and neurosurgery outpatient clinics

The results of Model Fitting Information and Goodness-Of-Fit showed the selected regression model to fit well and to describe the dependent variable. Pseudo R-square is high in the selected regression model

(Nagelkerke 0.485), which shows that the selected independent variables describe well the dependent variable (Pearson sig. 0.000, Deviance sig. 1.000). The test of parallel lines shows that null hypothesis can be rejected as p-value is >0.05, thus the model fits well and independent variables describe the dependent variable well. These results from analysis of goodness-of-fit and Pseudo R-square are shown in Appendix G: Model fitting information of ordinal logistic regression model.

The findings of ordinal logistic regression model showed that the main factors affecting the final significance of ENMG tests were as follows (put in order according to the size of estimate), see Table 3:

1. Prior evaluated insignificance was associated with lower final significance (-3.934, p-value 0.000)
2. Neurophysiologist was uncertain or did not have working diagnosis increased the significance (1.852, p-value 0.054)
3. ENMG finding being unsurprising lowered the final significance (-1.516, p-value 0.000)
4. Referring quarter was physiatry outpatient clinic or rehabilitation center it lowered the final significance (-1.374, p-value 0.009)
5. Neurophysiology intern performing the test lowered the final significance (-1.197, p-value 0.000)
6. ENMG finding being normal was associated with lower final significance (-.841, p-value 0.019)

The effects of independent variables are analyzed in more detail in next sub-chapters with chi-square test.

4.2.1 Prior evaluation of ENMG test's significance

According to ordinal logistic regression model prior evaluation of low significance was strongly associated with low final significance (estimate for low prior significance -3.934, p-value 0.000). This means that if neurophysiologists evaluated ENMG to be less significant based on a test referral, the test performed showed to be less significant according to final evaluation as well. This indicates that neurophysiologists seem to be able to evaluate significance of ENMG tests by reviewing ENMG test referrals. Even analysis of separate survey answers showed that 73.7 % (n=14) of insignificant ENMG tests were evaluated beforehand to be only slightly significant according to test referral. In addition, 48.8 % (n=42) of very significant ENMG tests were evaluated in advance to be very significant based on test referrals. Chi-square statistics test how the distributions of categorical, non-parametric variables differ from one another between the selected groups. The difference between groups of different significance levels is statistically significant (χ^2 118.02, p-value 0.000) (Table 4 below), which supports the finding from ordinal logistic regression model that prior evaluated significance is strongly associated with final significance.

Table 4 Effect of prior significance on final significance, chi-square test

	Final significance							
Prior significance	Insignificant	Slightly significant	Somewhat significant	Significant (needed for documentation)	**Very significant**	**Chi square -test value**	**p-value**	**sig.**
Slightly significant	73.7% (n=14)	80% (n=4)	25.3% (n=38)	6.3% (n=1)	5.8% (n=5)	118.0	**0.000**	**<0.05**
Somewhat significant	15.8% (n=3)	20% (n=1)	70% (n=105)	87.5% (n=14)	45.3% (n=39)			
Very significant	10.5% (n=2)	0% (n=0)	4.7% (n=7)	6.3% (n=1)	48.8% (n=42)			

Furthermore, the prior evaluation of ENMG tests' significance shows to have strong positive correlation with final evaluation of ENMG tests' significance (Spearman correlation 0.540, p-value 0.000) (Table 5 below). All these findings confirm the phenomenon that neurophysiologists were able to evaluate ENMG test's overall significance based on test referral.

Table 5 Spearman's rho between selected variables. (. Correlation is significant at the 0.01 level (2-tailed).)**

Correlations

			Final significance	Prior significance	Referral query	Certainty of working diagnosis	Evaluation of extent	Surprising findings	Extent of test
Spearman's rho	Final significance	Correlation Coefficient	1.000	.540**	.210**	-.006	.002	.363**	.050
		Sig. (2-tailed)	.	.000	.001	.921	.977	.000	.422
		N	263	263	263	263	263	263	263
	Prior significance	Correlation Coefficient	.540**	1.000	.308**	.011	.001	.279**	.047
		Sig. (2-tailed)	.000	.	.000	.855	.984	.000	.443
		N	263	263	263	263	263	263	263
	Referral query	Correlation Coefficient	.210**	.308**	1.000	.224**	.398**	.091	-.048
		Sig. (2-tailed)	.001	.000	.	.000	.000	.142	.437
		N	263	263	263	263	263	263	263
	Certainty of working diagnosis	Correlation Coefficient	-.006	.011	.224**	1.000	.093	-.112	-.035
		Sig. (2-tailed)	.921	.855	.000	.	.133	.069	.575
		N	263	263	263	263	263	263	263
	Evaluation of extent	Correlation Coefficient	.002	.001	.398**	.093	1.000	-.048	-.191**
		Sig. (2-tailed)	.977	.984	.000	.133	.	.439	.002
		N	263	263	263	263	263	263	263
	Surprising findings	Correlation Coefficient	.363**	.279**	.091	-.112	-.048	1.000	.188**
		Sig. (2-tailed)	.000	.000	.142	.069	.439	.	.002
		N	263	263	263	263	263	263	263
	Extent of test	Correlation Coefficient	.050	.047	-.048	-.035	-.191**	.188**	1.000
		Sig. (2-tailed)	.422	.443	.437	.575	.002	.002	.
		N	263	263	263	263	263	263	263

4.2.2 Neurophysiologists' certainty of working diagnosis

Regression model showed if neurophysiologist did not have any working diagnosis, it increased the final significance of ENMG test (estimate for no working diagnosis is 1.852, p-value 0.05). This phenomenon could be related to the effect that a surprising ENMG test finding is evaluated to be more significant. If neurophysiologist do not have any suggestions for working diagnosis, ENMG test most likely provides new information and the ENMG finding can be considered to be surprising. However, Spearman correlation shows that there is no correlation between certainty of working diagnosis and surprising finding (Spearman correlation -0.112, p-value 0.069).

Furthermore, chi-square test analysis of neurophysiologists' answers supported the finding from regression model that low certainty of working diagnosis increased ENMG test's final significance (χ^2 42.38, p-value 0.001) (Table 6 below). The more certain neurophysiologist was about the working diagnosis, the less significant ENMG test was performed. Since, 50 % (n=33) of ENMG tests with no working diagnoses were evaluated as very significant, while only 22.2 % (n=6) of ENMG tests with very certain working diagnoses resulted in very significant ENMG test, it seems that low certainty of proposed working diagnosis results in more significant ENMG test.

Table 6 Effect of certainty of working diagnosis on final significance, chi-square test

Certainty of working diagnosis	**Insignificant**	**Slightly significant**	**Somewhat significant**	**Significant (needed for documentation)**	**Very significant**	**Chi square - test value**	**p-value**	**sig.**
No working diagnosis	10% (n=1)	0% (n=0)	40% (n=4)	0% (n=0)	50% (n=5)	42.38	**0.001**	**<0.05**
Uncertain	2.7% (n=3)	3.6% (n=4)	63.6% (n=70)	0% (n=0)	30% (n=33)			
Almost certain	7% (n=9)	0.8% (n=1)	51.9% (n=67)	7.8% (n=10)	32.6% (n=42)			
Very certain	22.2% (n=6)	0% (n=)	33.3% (n=9)	22.2% (n=6)	22.2% (n=6)			

However, the quality of referral query affected neurophysiologist's certainty of working diagnosis (Spearman correlation 0.224, p-value 0.000). The poorer referral query, the more uncertain neurophysiologist was about the working diagnosis. Furthermore, 82 % (n=23) of ENMG tests with very certain working diagnoses had a well-defined and relevant referral query, while 30 % (n=33) of ENMG tests with uncertain working diagnoses had only somewhat defined and unclear referral query (χ^2 28.99, p-value 0.012) (Table 7 below).

Table 7 Effect of quality of referral query on the certainty of working diagnosis, chi-square test

The quality of referral query / Certainty of working diagnosis	Referral query missing	Unclear and poorly defined	Somewhat defined, but unclear	Well defined and clear, but irrelevant	Well defined, clear and relevant	Chi square -test value	p-value	sig.
No working diagnosis	9 % (n=1)	9% (n=1)	45% (n=5)	0% (n=0)	36% (n=4)	28.99	**0.012**	**<0.05**
Uncertain	0% (n=0)	10% (n=11)	30% (n=33)	6% (n=7)	54% (n=60)			
Almost certain	1% (n=1)	5% (n=6)	18% (n=24)	7% (n=9)	70% (n=92)			
Very certain	4% (n=1)	7% (n=2)	7% (n=2)	0% (n=0)	82% (n=23)			

4.2.3 Surprising ENMG finding and the significance of normal finding

Surprising ENMG finding was evaluated as significant ENMG test, whilst unsurprising finding as less significant. This effect was noticed in ordinal regression model as well; unsurprising ENMG test finding resulted in less significant ENMG test (estimate for unsurprising ENMG test result -1.516, p-value 0.000). Surprising ENMG finding correlated positively with both prior and final evaluation of ENMG tests' significance. Surprising finding correlated more positively with final evaluation of ENMG tests' significance (Spearman correlation 0.36, p-value 0.000) than with prior evaluation (Spearman correlation 0.28, p-value 0.000). These indicated that the more surprising ENMG test result, the more significant ENMG test was.

In addition, chi-square test of neurophysiologists' answers supported the finding, while 2.6 % (n=2) of insignificant ENMG tests gave surprising finding, compared to very significant ENMG tests of which 62.8 % (n=49) gave surprising finding (χ^2 31.87, p-value 0.000) (Table 8 below).

Table 8 Effect of surprising finding on final significance, chi-square test

Final significance / Surprising finding	**Insignificant and slightly significant**	**Somewhat significant**	**Significant and very significant**	**Chi square - test value**	**p-value**	**sig.**
Not surprising finding	11.2% (n=22)	61.9% (n=122)	26.9% (n=53)	31.87	**0.000**	**<0.05**
Surprising finding	2.6% (n=2)	34.6% (n=27)	62.8% (n=49)			

In addition to unsurprising finding, ENMG test giving a zero result, referring to normal finding, was evaluated less significant than test providing an abnormal finding. Since, 37.6 % (n=71) of abnormal ENMG tests were evaluated as very significant compared to 17.2 % (n=15) of normal findings evaluated as very significant (χ^2 12.00, p-value 0.02) (Table 9). Regression model supported this finding, as normal finding decreased the significance of ENMG test (estimate for normal test result -0.841, p-value 0.019).

Table 9 Effect of normal finding on final significance, chi-square test

Final significance / Normal or abnormal finding	**Insignificant**	**Slightly significant**	**Somewhat significant**	**Significant (needed for documentation)**	**Very significant**	**Chi square test value**	**p-value**	**sig.**
Abnormal	5.8% (n=11)	1.6% (n=3)	50.3% (n=95)	4.8% (n=9)	37.6% (n=71)	12	**0.017**	**<0.05**
Normal	9.2% (n=8)	2.3% (n=2)	63.2% (n=55)	8% (n=7)	17.2% (n=15)			

4.2.4 Difference between neurophysiology specialist and intern

According to survey answers from neurophysiologists, final evaluation of ENMG test correlates with neurophysiologist's degree of specialty; the lower the degree of specialty of neurophysiologist, the less significant ENMG test was (Pearson correlation 0.259, p-value 0.000) (Appendix H: Pearson correlation between ordinal logistic regression model variables). Furthermore, a more detailed analysis of survey answers showed that greater percentage of ENMG tests performed by neurophysiology specialists were evaluated to be very significant while ENMG tests performed by neurophysiology intern were evaluated to be less significant; 40.5 % (n=66) of ENMG tests performed by specialist were evaluated as very significant, while 16.8 % (n=20) of interns' ENMG tests were evaluated as very significant (χ^2 42.30, p-value 0.000) (Table 10). A majority of ENMG tests conducted by interns was evaluated by themselves to be only somewhat significant (73.9 % n=88). The same effect was noticed in ordinal logistic regression model, where it showed that interns conducted less significant ENMG tests than specialists (estimate -1.197, p-value 0.000).

Table 10 Effect of the level of specialty of neurophysiologist on final significance

Neurophysiologist who performed ENMG test / Final significance	**Specialist**	**Intern**	**Chi-square test - value**	**p-value**	**sig.**
Insignificant	5.5% (n=9)	8.4% (n=10)	42.30	**0.000**	**<0.05**
Slightly significant	3.1% (n=5)	0% (n=0)			
Somewhat significant	38% (n=62)	73.9% (n=88)			
Significant (needed for documentation)	9.2% (n=15)	0.8% (n=1)			
Very significant	40.5% (n=66)	16.8% (n=20)			

A reason for such difference between specialists and interns is not described by the different patient mix, because the distribution of ENMG tests requested by different referring quarters was the same between the two groups (χ^2 4.33, p-value 0.362) (Table 11 below). However, the difference may be due to the different opinions about what test should be considered as significant or insignificant, because these answers were based on their own evaluation.

Table 11 Distribution of different referring quarters between interns and specialists

Neurophysiologist who performed ENMG test / Referring quarter	**Specialist**	**Intern**	**Chi-square test -value**	**p-value**	**sig.**
Hand surgery and surgery outpatient clinic and orthopedic outpatient clinic	26.4% (n=43)	20.2% (n=24)	4.33	0.362	>0.05
Physiatry outpatient clinic and rehabilitation center	14.7% (n=24)	10.1% (n=12)			
Other outpatient clinics	16% (n=26)	16% (n=19)			
Neurology outpatient clinic and neurosurgery outpatient clinic	19.6% (n=32)	21.8% (n=26)			
Primary care	23.3% (n=38)	31.9% (n=38)			

4.2.5 Other factors affecting the significance of ENMG test

In addition to ordinal regression model's statistically significant independent variables, through chi-square test analysis also other factors were identified to affect ENMG tests' significance. Especially, the quality of referral query seemed to affect ENMG tests' final significance. Although, according to regression model referral query was not statistically significant to affect final significance of an ENMG. Nonetheless, according to separate analysis of survey answers the poorer the referral query, the less significant ENMG was performed. Only 37.5 % (n=9) of ENMG tests with poorly defined, missing or unclear referral query resulted in very significant ENMG test. In contrast, 74.4 % (n=64) of ENMG tests with a well-defined, relevant and clear referral query resulted in very significant ENMG test. In other words, the significance of ENMG tests varied according to the quality of referral query (χ^2 16.96, p-value 0.031) (Table 12 below). Furthermore, there seemed to be a positive statistically significant correlation between the quality of referral query and the final significance of ENMG test (Spearman correlation 0.210, p-value 0.001). Similar and stronger correlation seemed to be between the quality of referral query and prior significance (Spearman correlation 0.308, p-value 0.000).

Table 12 Effect of the quality of referral query on final significance, chi-square test

Final significance / Quality of referral query	**Insignificant**	**Slightly significant**	**Somewhat significant**	**Significant (needed for documentation)**	**Very significant**	**Chi square -test value**	**p-value**	**sig.**
Referral query missing or unclear and poorly defined	4.2% (n=1)	16.7% (n=4)	29.2% (n=7)	12.5% (n=3)	37.5% (n=9)	16.95	**0.031**	**<0.05**
Somewhat defined, but unclear or well-defined, but irrelevant	1.2% (n=2)	8.4% (n=14)	23.5% (n=39)	5.4% (n=9)	61.4% (n=102)			
Well defined, clear and relevant	0% (n=0)	2.3% (n=2)	18.6% (n=16)	4.7% (n=4)	74.4% (n=64)			

The extent of performed ENMG test was not affected by the quality of test referral query. The assumption was that the poorer the test referral query was, the more extensive ENMG test was performed. This assumption was based on the fact that if a test referral query was unfocused and poorly defied it would result in more extensive ENMG test, because a neurophysiologist would need to do screening and search for the reason for patient's symptoms. Therefore, the extent did not affect the final significance of ENMG test and the poorly defined test referral did not result in ENMG test to be more extensive (χ^2 4.15, p-value 0.386) (Table 13 and Table 14 below).

Table 13 The effect of the quality of test referral on the extent of ENMG test, chi-square test

The quality of referral query / Extent of ENMG test	**Referral query missing or unclear and poorly defined**	**Somewhat defined, but unclear or well-defined, but irrelevant**	**Well defined and relevant**	**All**	**Chi square - test value**	**p-value**	**sig.**
ENMG-1	0% (n=0)	3.8% (n=3)	2.3% (n=4)	2.5% (n=7)	4.15	0.386	>0.05
ENMG-2	52.2% (n=12)	68.8% (n=55)	65% (n=115)	65% (n=182)	2.99 *(between ENMG-2 and ENMG-3)*	0.22 *(between ENMG-2 and ENMG-3)*	
ENMG-3	47.8% (n=11)	27.5% (n=22)	32.8% (n=58)	32.5% (n=91)			

Table 14 The effect of the extent of ENMG test on the final significance of ENMG test, chi-square test

Final significance / The extent of ENMG test	**Insignificant or slightly significant**	**Somewhat significant**	**Significant or very significant**	**Chi square - test value**	**p-value**	**sig.**
ENMG-1	14.3% (n=1)	28.6% (n=2)	57.1% (n=4)	3.84	0.428	>0.05
ENMG-2	8.5% (n=15)	58% (n=102)	33.5% (n=59)			
ENMG-3	8.8% (n=8)	49.5% (n=45)	41.8% (n=38)			

In survey neurophysiologists were asked to list the main reasons for ENMG test's insignificance. The major reason was self-evident diagnosis, which accounted 42 % (n=10) of all insignificant or slightly significant tests. The other two specific reasons were unnecessary control test 20.8% (n=5), and too extensive and unspecific referral query 20.8 % (n=5). Other reasons included reasons such as normal finding, patient's symptoms had disappeared, same ENMG finding found previously which had not reached referring physician, and the fact that patient's symptoms did not relate to nervous system dysfunction (Table 15 below).

Table 15 Reasons presented for insignificant and slightly significant ENMG test

Final significance / reasons for ENMG tests' insignificance	Insignificant or slightly significant
Other reason	16.7% (n=4)
Unnecessary control test	20.8% (n=5)
Self-evident diagnosis	41.7% (n=10)
Too broad and unspecific referral query, which caused ENMG finding to be irrelevant for patient's symptoms	20.8% (n=5)
Other reasons included:	
normal finding	
symptoms had disappeared	
same ENMG test done previously but information had not reached referring doctor	
patient's symptoms did not relate to nervous system dysfunction plus normal ENMG test result	

To summarize the findings of chi-square test analysis of neurophysiologists' answers, the main factors contributing to insignificant ENMG test were prior evaluation of low significance, normal finding, unsurprising finding, poor quality of referral query, if neurophysiologist intern performed the test, and if neurophysiologist was certain about the working diagnosis.

In addition to aforementioned characteristics of insignificant ENMG test, the findings of ordinal logistic regression model showed that the referring quarter affected ENMG tests' significance. Due to this a more detailed analysis was made to identify the differences between referring quarters. The results from these analyses are presented separately in the next chapter.

4.3 The differences between referring quarters

This chapter explores the differences between primary care and secondary care from the ENMG test requestor perspective. Secondary care is analyzed in more detail in order to explore the differences between different outpatient clinics. Lastly, this chapter reveals the findings of where to focus demand management in future in order to decrease the number of inappropriate ENMG test requests. The analysis in this chapter is based on surveys performed to neurophysiologists.

According to ordinal logistic regression model physiatry outpatient clinics and rehabilitation center requested less significant ENMG tests compared to other referring quarters including primary care (estimate for physiatry outpatient clinic and rehabilitation center -1.374, p-value 0.009) (Table 2 above). However, a more detailed analysis of neurophysiologists' answers to survey showed a great difference between primary care and secondary care requests. Primary care orders less significant ENMG tests compared to secondary care according to neurophysiologists' final evaluation of ENMG tests' significance. The overall percentage of insignificant tests was 6.7 % and only slightly significant ENMG tests was 1.8 %, while the percentage of very significant ENMG tests was 30.5 % (n=86). According to final evaluation, 19.2 % (n=15) of ENMG tests ordered by primary care physicians were evaluated as very significant while similarly significant ENMG tests accounted 34.8% (n=71) of ENMG tests requested by secondary care. There was statistically significant difference between primary care and secondary care at the significance level of 90 % (χ^2 9.37, p-value 0.09) (Table 16 below).

Primary care ordering less significant ENMG tests is explained with various differences compared to secondary care concerning ENMG test requesting.

Table 16 Differences between primary and secondary care ENMG test request, chi-square test

	All physicians	Primary care	Secondary care	Chi square -test	p-value	sig.
Neurophysiologist final evaluation of ENMG tests' significance	*100 % (n=282)*	*100% (n=78)*	*100% (n=204)*			
Insignificant	6.7% (n=19)	7.7% (n=6)	6.4% (n=13)	9.37	0.090	>0.05, **<0.1**
Slightly significant	1.8% (n=5)	0.0% (n =0)	2.5% (n=5)			
Somewhat significant	53.2% (n=150)	59.0% (n=46)	51% (n=104)			
Significant (needed for documentation)	5.7% (n=16)	9.0% (n=7)	4.4% (n=9)			
Very significant	30.5% (n=86)	19.2% (n=15)	34.8% (n=71)			
No evaluation	2.1% (n=6)	5.1% (n=4)	1% (n=2)			
Prior significance of ENMG test	*100 % (n=282)*	*100 % (n=78)*	*100 % (n=204)*			
Insignificant	0% (n=0)	0% (n=0)	0% (n=0)	10.43	0.000	**<0.05**
Slightly significant	22% (n=62)	24.4% (n=19)	21.1% (n=43)			
Somewhat significant	59.6% (n=168)	69.2% (n=54)	55.9% (n=114)			
Very significant	18.4% (n=52)	6.4% (n=5)	23% (n=47)			
ENMG test results	*100 % (n=282)*	*100 % (n=78)*	*100 % (n=204)*			
Abnormal results	67% (n=189)	55.1% (n=43)	71.6% (n=146)	6.90	0.009	**<0.05**
Normal results	33% (n=93)	44.9% (n=35)	28.4% (n=58)			
Surprising ENMG test findings	*100% (n=275)*	*100% (n=73)*	*100% (n=202)*			
Unsurprising abnormal results	45.8% (n=126)	42.5% (n=31)	47% (n=95)	11.64	0.003	**<0.05**
Unsurprising normal results	25.8% (n=71)	38.4% (n=28)	21.3% (n=43)			
Surprising abnormal results	21.5% (n=59)	11% (n=8)	25.2% (n=51)			
surprising normal results	6.9% (n=19)	8.2% (n=6)	6.4% (n=13)			

Table 17 Differences between primary and secondary care's ENMG test requesting, chi-square test

	All physicians	Primary care	Secondary care	Chi square -test	p-value	sig.
Referral diagnosis	*100 % (n=282)*	*100 % (n=78)*	*100 % (n=204)*			
No referral diagnosis	13.8% (n=39)	10.5% (n=8)	15% (n=31)	17.08	0.003	**<0.05**
Radiculopathy (unspecific)	12.8% (n=36)	10.5% (n=8)	13.6% (n=28)			
Polyneuropathy	14.9% (n=42)	21.1% (n=16)	12.6% (n=26)			
Lower limb radiculopathy	11.7% (n=33)	13.2% (n=10)	11.2% (n=23)			
CTS	9.6% (n=27)	14.5% (n=11)	7.8% (n=16)			
Too general diagnosis	6% (n=17)	11.8% (n=9)	3.9% (n=8)			
Ulnar nerve neuropathy	5% (n=14)	1.3% (n=1)	6.3% (n=13)			
Other upper limb neuropathy	2.5% (n=7)	0% (n=)	3.4% (n=7)			
Other pathologies	18.8% (n=53)	13.2% (n=10)	20.9% (n=43)			
% of times referral physician was right about referral diagnosis	*100 % (n=varies)*	*100 % (n=varies)*	*100 % (n=varies)*			
Radiculopathy (unspecific)	52.8% (n=19)	50% (n=4)	53.6% (n=15)			
Lower limb radiculopathy	55.6% (n=15)	16.7% (n=1)	66.7% (n=14)			
CTS	63% (n=17)	63.6% (n=7)	62.5% (n=10)			
Polyneuropathy	33.3% (n=14)	23.5% (n=4)	40% (n=10)			
Ulnar nerve neuropathy	71.4% (n=10)	100% (n=1)	69.2% (n=9)			
Upper limb radiculopathy	50% (n=5)	50% (n=1)	50% (n=4)			
Agreement of referral diagnosis with ENMG finding	*100 % (n=242)*	*100 % (n=69)*	*100 % (n=173)*			
No	61.7% (n=150)	72.5% (n=50)	57.5% (n=100)	4.70	0.02	**<0.05**
Yes	38.3% (n=93)	27.5% (n=19)	42.5% (n=74)			
Agreement of working diagnosis with ENMG finding	*100 % (n=264)*	*100 % (n=76)*	*100 % (n=188)*			
No	56.4% (n=149)	68.4% (n=52)	51.6% (n=97)	6.23	0.01	**<0.05**
Yes	43.6% (n=115)	31.6% (n=24)	48.4% (n=91)			

Table 18 Differences between primary and secondary care's ENMG test requesting, chi-square test

	All physicians	**Primary care**	**Secondary care**	**Chi square -test**	**p-value**	**sig.**
Agreement of referral diagnosis with working diagnosis	*100% (n=235)*	*100% (n=69)*	*100% (n=166)*			
No	34% (n=80)	42% (n=29)	30.7% (n=51)	2.77	0.10	>0.05
Yes	66% (n=155)	58% (n=40)	69.3% (n=115)			
Certainty of working diagnosis	*100 % (n=282)*	*100% (n=78)*	*100% (n=204)*			
Not certain	39.4% (n=111)	37.2% (n=29)	40.2% (n=82)	1.18	0.72	>0.05
Almost certain	46.8% (n=132)	51.3% (n=40)	45.1% (n=92)			
Very certain	9.9% (n=28)	9% (n=7)	10.3% (n=21)			
Don't have clinical diagnosis	3.9% (n=11)	2.6% (n=2)	4.4% (n=9)			
The quality of referral query	*100 % (n=282)*	*100 % (n=78)*	*100 % (n=204)*	*Referral query missing or unclear vs. referral query well-defined*		
Referral query missing	1.1% (n=3)	1.3% (n=1)	1% (n=2)	2.93	0.25	>0.05
Unclear and poorly defined referral query	7.1% (n=20)	9% (n=7)	6.4% (n=13)			
Somewhat defined query, but still unclear	22.7% (n=64)	28.2% (n=22)	20.6% (n=42)			
Well defined and clear query, but still irrelevant	5.7% (n=16)	5.1% (n=4)	5.9% (n=12)			
Well defined, clear and relevant query	63.5% (n=179)	56.4% (n=44)	66.2% (n=135)			
Evaluation of ordered extent of ENMG test	*100 % (n=282)*	*100 % (n=78)*	*100 % (n=204)*			
Appropriate	85.8% (n=242)	89.7% (n=70)	84.3% (n=172)	0.29	0.59	>0.05
Too broad	8.5% (n=24)	9% (n=7)	8.3% (n=17)			
Too narrow	1.8% (n=5)	0% (n=)	2.5% (n=5)			
ENMG is wrong test for examining patient	0.4% (n=1)	0% (n=)	0.5% (n=1)			
Not evaluated	3.5% (n=10)	1.3% (n=1)	4.4% (n=9)			

First, according to neurophysiologists' prior evaluation, primary care requests less significant ENMG test than secondary care. Since, 23 % (n=47) of secondary care requested ENMG tests were considered to be very significant by evaluating the referrals while only 6.4 % (n=5) of primary care ENMG test referrals were considered to provide very significant ENMG test (χ^2 10.43, p-value 0.000) (Table 16 above).

Secondly, primary care physicians order more normal result tests than secondary care physicians. Of ENMG tests ordered by primary care, 44.9 % (n=35) gave normal finding compared to 28.4 % (n=58) of ENMG tests requested by secondary care physicians (χ^2 6.90, p-value 0.010 (Table 16 above).

This difference is not due to primary care physicians' greater number of exclusion diagnoses, because only 3.9 % (n=3) of primary care physicians ordered ENMG tests were meant to exclude a diagnosis compared to 10.2 % (n=21) of secondary care physicians ordered ENMG tests to exclude a diagnosis.

Thirdly, secondary care physicians requested ENMG tests that provided more surprising findings than primary care physicians. Especially, when comparing the number of unsurprising normal findings, primary care physicians order more of such tests. Since, 21.3 % (n=43) of secondary care physicians requested ENMG tests provided unsurprising normal finding compared to 38.4 % (n=28) of primary care physicians respectively. Neurophysiologists' opinion about self-evident zero results is that these kind of tests should be reduced, because examining patients who assumedly have no nervous system dysfunction, is waste of resources and causes unnecessary discomfort to a patient.

In contrast, the number of surprising abnormal findings was much greater among ENMG tests ordered by secondary care physicians. The number of ENMG tests gave surprising abnormal result was 11% (n=8) in primary care physicians compared to 25% (n=51) in secondary care physicians ordered tests. There is a statistically significant difference between secondary care and primary care in ordering of surprising normal and abnormal result ENMG tests (χ^2 11.64, p-value 0.000) (Table 16 above).

Lastly, the referral diagnoses suggested by referring physician differs between secondary care and primary care. The two most common referral diagnoses suggested by primary care was polyneuropathy (21.1%, 16/78) and carpal tunnel syndrome (CTS) (14.5 %, 11/78). However, primary care physicians included unfocused referral queries in 11.8% (9/78) of referrals, which was the third most common category of referral queries. In secondary care referrals the most common referral diagnoses were much more varied; other pathologies such as amyotrophic lateral sclerosis (ALS) and spinal stenosis (20.9 %, 43/204), no referral diagnosis (15%, 31/204), radiculopathy (unspecific) (13.6%, 28/204) and polyneuropathy (12.6 %, 26/204). There was a statistically significant difference between primary care and secondary care considering the referral diagnoses (χ^2 17.08, p-value 0.000) (Table 17 above).

Generally, secondary care physicians were more often correct with their referral diagnosis than primary care physicians. Agreement between referral diagnosis and ENMG finding was higher with secondary care physicians than with primary care physicians; 42.5 % (n=74) of secondary care physicians' and 27.5 % (n=19) of primary care physicians' referral diagnoses agreed with the final ENMG finding, There was a statistically significant difference between secondary care and primary care concerning the agreement of referral diagnosis and ENMG finding (χ^2 4.70, p-value 0.020) (Table 17 above).

Furthermore, these two referring quarters differ in referral diagnoses which they manage to confirm with ENMG test. Primary care physicians were most commonly correct with CTS diagnosis, namely in 63.6% (n=7) of times when CTS was suggested as referral diagnosis. Secondary care physicians had a correct referral diagnosis with multiple diagnoses such as ulnar nerve neuropathy (69.2 % of times when ulnar nerve was suggested, n=9), lower limb radiculopathy (66.7 % n=14), CTS (62.5 % n=10) and radiculopathy (53.6 % n=15) (Table 17 above).

The agreement between working diagnosis and final ENMG finding was slightly higher in secondary care patients than primary care patients. Altogether 48.4 % (n=91) of working diagnoses in secondary care patients compared to 31.6 % (n=24) of working diagnoses in primary care patients agreed with the final ENMG (χ^2 6.23, p-value 0.010) (Table 17 above). However, the percentage of agreement between referral diagnosis and working diagnosis did not differ between secondary and primary care (χ^2 2.77, p-value 0.100) (Table 18 above).

Although there are a lot of differences between secondary and primary care as referring quarters, they showed similar results in a couple of areas of requesting ENMG tests. First, neurophysiologist's certainty of used working diagnoses were similar between secondary care and primary care patients (χ^2 1.18, p-value 0.720) (Table 18 above). Second, both primary and secondary care physicians requested the same number of *i)* too extensive, *ii)* too narrow and *iii)* appropriate extent ENMG tests (χ^2 0.29, p-value 0.590) (Table 18 above). Lastly, the quality of referral query was similar between these two referring quarters (χ^2 2.93, p-value 0.25) (Table 18 above). 66.2 % (n=135) of secondary care referrals and 56.4 % (n=44) of primary care referrals included a well-defined, relevant and clear referral query. The overall percentage of inappropriate ENMG test referrals was 30.9 % (87/282) of all ENMG test referrals (Table 18 above). The percentage varied depending on the referring quarter, being 38.5 % among primary care physicians and 28 % among secondary care physicians.

In ordinal logistic regression model, physiatry outpatient clinic and rehabilitation center showed to request less significant ENMG tests than other referring quarters. In a separate analysis of neurophysiologists' answers, physiatry outpatient clinic and rehabilitation center seemed to have similar test requesting characteristics as primary care. The final significance of ENMG tests requested by physiatry outpatient clinic and rehabilitation center differed slightly although not significantly from other referring polyclinics; 19.4 % (n=7) of ENMG tests ordered by physiatry outpatient clinic and rehabilitation center were evaluated as very significant, while the percentage of very significant ENMG tests requested by all other outpatient clinics was 28.4 % – 44.4 % of all ENMG tests ordered (Table 19 below). This difference is not statistically significant when comparing all secondary care referring quarters with each other (χ^2 16.15, p-value 0.33), but comparing physiatry outpatient clinic and rehabilitation center alone to all other outpatient clinics the difference is statistically significant. Especially, neurology and neurosurgery outpatient clinics requested more very significant ENMG tests than physiatry outpatient clinic and rehabilitation center.

Table 19 Differences within secondary care referring units, chi-square test

	All referring quarters	**Hand surgery and general surgery and orthopedic outpatient clinic**	**Physiatry outpatient clinic and rehabilitation center**	**Other outpatient clinics**	**Neurology and neurosurgery outpatient clinics**	**Primary care**	**Chi square -test**	**p-value**	**sig.**
Neurophysiologist final evaluation of ENMG tests' significance	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*	*100% (n=76)*			
Insignificant	6.7% (n=19)	7.5% (n=5)	8.3% (n=3)	6.7% (n=3)	3.4% (n=2)	7.9% (n=6)	28.48	0.065	>0.05
Slightly significant	1.8% (n=5)	3% (n=2)	8.3% (n=3)	0% (n=0)	0% (n=0)	0% (n=0)			
Somewhat significant	53.2% (n=150)	52.2% (n=35)	58.3% (n=21)	44.4% (n=20)	48.3% (n=28)	60.5% (n=46)			
Significant (needed for documentation)	5.7% (n=16)	6% (n=4)	5.6% (n=2)	4.4% (n=2)	3.4% (n=2)	7.9% (n=6)			
Very significant	30.5% (n=86)	28.4% (n=19)	19.4% (n=7)	44.4% (n=20)	44.8% (n=26)	18.4% (n=14)			
No evaluation	2.1% (n=6)	3% (n=2)	0% (n=0)	0% (n=0)	0% (n=0)	5.3% (n=4)			

The second difference between physiatry outpatient clinic and rehabilitation center and all other outpatient clinics is the number of too extensive ENMG tests requested and performed. 19.4 % (n=7) of ENMG tests requested by physiatry outpatient clinic and rehabilitation center were evaluated to be too extensive compared to 3 % - 10.3 % of ENMG tests ordered by all other outpatient clinics (χ^2 6.96, p-value 0.073). This difference between referring quarters can be considered as statistically significant at the significance level of 90 % (Table 20 below). The greatest difference was between physiatry outpatient clinic and rehabilitation center and surgery and orthopedic outpatient clinics.

Table 20 Differences within secondary care referring units, chi-square test

	All referring quarters	**Hand surgery and general surgery and orthopedic outpatient clinic**	**Physiatric outpatient clinic and rehabilitation center**	**Other outpatient clinics**	**Neurology and neurosurgery outpatient clinics**	**Primary care**	**Chi square - test**	**p-value**	**sig.**
Extent of ENMG test	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*				
Appropriate	85.8% (n=242)	88.1% (n=59)	77.8% (n=28)	91.1% (n=41)	79.3% (n=46)		6.96	0.073	>0.05, **<0.10**
Too broad	8.5% (n=24)	3% (n=2)	19.4% (n=7)	4.4% (n=2)	10.3% (n=6)				
Too narrow	1.8% (n=5)	4.5% (n=3)	2.8% (n=1)	0% (n=0)	1.7% (n=1)				
ENMG is wrong test for examining patient	0.4% (n=1)	0% (n=0)	0% (n=0)	2.2% (n=1)	0% (n=0)				
Not evaluated	3.5% (n=10)	4.5% (n=3)	0% (n=0)	2.2% (n=1)	8.6% (n=5)				
Agreement of referral diagnosis with ENMG finding	*100% (n=226)*	*100% (n=51)*	*100% (n=32)*	*100% (n=36)*	*100% (n=48)*				
No	58.8% (n=133)	41.2% (n=21)	68.8% (n=22)	50% (n=18)	64.6% (n=31)		8.53	0.036	**<0.05**
Yes	41.2% (n=93)	58.8% (n=30)	31.3% (n=10)	50% (n=18)	35.4% (n=17)				
% of exclusion diagnoses of 'no agreement' between referral diagnosis and ENMG finding	10.1% (n=19)	13.5% (n=5)	7.7% (n=2)	18.5% (n=5)	12.2% (n=5)		7.49	0.28	>0.05

Lastly, the agreement of referral diagnosis and final ENMG finding was lower in physiatry outpatient clinic than in all other outpatient clinics. Physiatry outpatient clinic's and rehabilitation center's referral diagnoses agreed with final ENMG finding in 31.3 % (n=10) of all referral diagnoses suggested compared to 35.4 % - 58.8 % of referral diagnoses suggested by all other outpatient clinics. The difference between these referring quarters concerning the diagnosis agreement is statistically significant (χ^2 8.53, p-value 0.04) (Table 20 above). Physiatry outpatient clinic and rehabilitation center did have only 7.7 % (n=2) of their referral diagnoses suggested for exclusion diagnosis purpose, while the same percentage varied between 12.2 – 18.5 % in all other outpatient clinics (Table 20). Therefore the low agreement between referral diagnosis and ENMG finding was not due to great number of exclusion diagnoses suggested by physiatry outpatient clinic and rehabilitation center. The greatest difference concerning the agreement of referral diagnosis and ENMG finding was between physiatry outpatient clinic and rehabilitation center and surgery and orthopedic outpatient clinics.

Table 21 Differences within secondary care referring units, chi-square test

	All referring quarters	Hand surgery and general surgery and orthopedic outpatient clinic	Physiatric outpatient clinic and rehabilitation center	Other outpatient clinics	Neurology and neurosurgery outpatient clinics	Primary care	Chi square - test	p-value	sig.
The quality of referral query	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*	*100% (n=76)*			
Referral query missing	1.1% (n=3)	3% (n=2)	0% (n=0)	0% (n=0)	0% (n=0)	1.3% (n=1)	10.26	0.247	>0.05
Unclear and poorly defined	7.1% (n=20)	6% (n=4)	8.3% (n=3)	8.9% (n=4)	3.4% (n=2)	9.2% (n=7)			
Somewhat defined, but still unclear	22.7% (n=64)	23.9% (n=16)	30.6% (n=11)	6.7% (n=3)	20.7% (n=12)	28.9% (n=22)			
Well defined and clear, but still irrelevant	5.7% (n=16)	3% (n=2)	8.3% (n=3)	8.9% (n=4)	5.2% (n=3)	5.3% (n=4)			
Well defined, clear and relevant	63.5% (n=179)	64.2% (n=43)	52.8% (n=19)	75.6% (n=34)	70.7% (n=41)	55.3% (n=42)			
ENMG test results	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*				
Abnormal results	68.4% (n=193)	71.6% (n=48)	66.7% (n=24)	73.3% (n=33)	75.9% (n=44)		0.98	0.807	>0.05
Normal results	31.6% (n=89)	28.4% (n=19)	33.3% (n=12)	26.7% (n=12)	24.1% (n=14)				
Unsurprising abnormal results	52.8% (n=149)	61.2% (n=41)	52.8% (n=19)	51.1% (n=23)	51.7% (n=30)		8.69	0.466	>0.05
Unsurprising normal results	25.9% (n=73)	23.9% (n=16)	27.8% (n=10)	20% (n=9)	17.2% (n=10)				
Surprising abnormal results	13.8% (n=39)	10.4% (n=7)	13.9% (n=5)	22.2% (n=10)	24.1% (n=14)				
Surprising normal results	5% (n=14)	1.5% (n=1)	5.6% (n=2)	6.7% (n=3)	6.9% (n=4)				

Table 22 Differences within secondary care referring units, chi-square test

	All referring quarters	Hand surgery and general surgery and orthopedic outpatient clinic	Physiatric outpatient clinic and rehabilitation center	Other outpatient clinics	Neurology and neurosurgery outpatient clinics	Primary care	Chi square - test	p-value	sig.
Certainty of working diagnosis	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*				
Not certain	39.4% (n=111)	32.8% (n=22)	44.4% (n=16)	48.9% (n=22)	37.9% (n=22)		10.85	0.286	>0.05
Almost certain	46.8% (n=132)	50.7% (n=34)	41.7% (n=15)	35.6% (n=16)	48.3% (n=28)				
Very certain	9.9% (n=28)	13.4% (n=9)	2.8% (n=1)	13.3% (n=6)	8.6% (n=5)				
Don't have clinical diagnosis	3.9% (n=11)	3% (n=2)	11.1% (n=4)	2.2% (n=1)	5.2% (n=3)				
Referral diagnosis	*100% (n=282)*	*100% (n=67)*	*100% (n=36)*	*100% (n=45)*	*100% (n=58)*				
No referral diagnosis	19.9% (n=56)	23.9% (n=16)	11.1% (n=4)	20% (n=9)	17.2% (n=10)		2.14	0.543	>0.05
Agreement of referral diagnosis with working diagnosis	*100% (n=217)*	*100% (n=51)*	*100% (n=27)*	*100% (n=36)*	*100% (n=45)*				
No	28.6% (n=62)	23.5% (n=12)	33.3% (n=9)	30.6% (n=11)	24.4% (n=11)		1.24	0.743	>0.05
Yes	71.4% (n=155)	76.5% (n=39)	66.7% (n=18)	69.4% (n=25)	75.6% (n=34)				

Physiatry outpatient clinic and rehabilitation center did not differ in other areas of test requesting from other referring quarters than above mentioned (Table 21, Table 22). Physiatry outpatient clinic and rehabilitation center referrals were evaluated to be similar quality as all other referring quarters' referrals. Same phenomenon was noticed with primary care referrals.

4.4 Comparison of referring physician's and neurophysiologist's opinions

In this chapter the opinion about ENMG tests' significance is explored and compared between referring physicians and neurophysiologists. In addition, this chapter presents the percentages of insignificant ENMG tests evaluated by both referring physicians and neurophysiologists.

Survey was conducted to all referring quarters electronically, however, after the first week of introducing the survey to the quarters, we noticed that it did not reach the referring units. Therefore a paper version of the same survey was formulated. However, this survey was performed only to the top 10 most ENMG test requesting referring units. We gathered 70 answers from the referring quarters. The majority of answers came from physiatrists (21.4 %), hand surgeons (30 %) and neurologists (28.6 %), while we got only few answers from primary care physicians (2.9 %). The work expertise of referring physicians who answered to the survey was high; 50 % had 5-10 years of working experience as doctor and 42.9 % had over 10 years of working experience (Table 23 below).

Table 23 the profile of answerers of a survey made to referring physicians

Physician typology	
Physiatrist	21.4% (n=15)
Hand surgeon	30% (n=21)
Orthopedist	1.4% (n=1)
Neurologist	28.6% (n=20)
Intern	14.3% (n=10)
General practitioner	2.9% (n=2)
Anesthesiologist	1.4% (n=1)

Experience	
1-5 years	7.1% (n=5)
5-10 years	50% (n=35)
over 10 years	42.9% (n=30)

Findings of this survey were interesting. According to referring physicians, 42.9 % (n=30) of all ENMG tests requested, did not change patients' treatment plan. 17.1 % (n=12) of ENMG tests changed already decided treatment plan, and 40 % (n=28) of ENMG tests changed treatment plan as the decision was meant to be made according to ENMG test results (Table 24 below). Of ENMG tests ordered, 35.7 % (n=25) excluded a diagnosis, while 44.3 % (n=31) confirmed a diagnosis suspicion. Only 10 % of ENMGs (n=7) provided new diagnosis alternative. 7.1 % (n=5) were ordered for control and the remaining 2.9 % (n=2) did not affect patient treatment plan because ENMG test result was in contrast to other diagnostic test results. (Table 24 below).

Table 24 Results of how ENMG test affected referring physicians' decision-making.

ENMG test result changed patient's treatment plan	
No	42.9% (n=30)
Yes	17.1% (n=12)
Treatment plan was meant to do according to ENMG test result	40% (n=28)
ENMG test affected patient diagnosis by:	
Excluded diagnosis	35.7% (n=25)
Confirmed diagnosis	44.3% (n=31)
Provided new diagnosis	10% (n=7)
Control test	7.1% (n=5)
Something else:	2.9% (n=2)
Clinical examination tells different than ENMG test results	
MRI test result tells different than ENMG test result	

There were only 33 patient cases of which the opinions of ENMG test significance was gathered both referring physicians and neurophysiologists. Because the sample size was small, no statistical significance analysis could be performed. However, opinions seemed to be mostly in line, and especially when comparing neurophysiologists' prior evaluation of significance and referring physicians' final evaluation of significance for patient diagnosing and treatment plan decision-making. The majority of the tests, namely 71.8 % (24/33) of all ENMG tests, were evaluated similarly by both parties, although there were a couple of cases which were dissimilarly evaluated. Altogether 63.7 % (n=21) of all ENMG tests, evaluated by both parties, was evaluated somewhat significant or very significant by both referring physicians and neurophysiologists (Table 25 below). Furthermore, 9.1 % (n=3) of these ENMG tests were evaluated insignificant or slightly significant by both parties (Table 25). However, the remaining 27.3 % (n=9) was evaluated dissimilarly. Nonetheless, neurophysiologist evaluated 67 % (n=6) of these dissimilarly evaluated ENMG tests to be either somewhat significant or very significant, thus neurophysiologist would have conducted these tests anyway.

Table 25 Comparison of referring physician's and neurophysiologist's opinion about ENMG test's significance, N=33. Referring physician’s evaluation of ENMG test’s significance for treatment plan decision-making.

		Neurophysiologist’s prior evaluation of the significance of ENMG test			
		Insignificant	Slightly significant	Somewhat significant	Very significant
Referral physician's evaluation of test's significance for treatment plan	Insignificant	0% (n=0)	6.1% (n=2)	0% (n=0)	0% (n=0)
	Slightly significant	0% (n=0)	3% (n=1)	12.1% (n=4)	6.1% (n=2)
	Somewhat significant	0% (n=0)	0% (n=0)	36.4% (n=12)	6.1% (n=2)
	Very significant	0% (n=0)	9.1% (n=3)	18.2% (n=6)	3% (n=1)

Referring physicians evaluated the significance from the diagnosing point of view and it was compared to neurophysiologists’ prior evaluation of significance. 66.7 % (n=22) of these were evaluated either somewhat significant or very significant by both parties (Table 26 below). 6.1 % (n=2) of ENMG tests were evaluated slightly significant by both parties (Table 26). 21.1 % (n=7) were evaluated dissimilarly, however 57 % (n=4) of dissimilarly evaluated ENMG tests were evaluated somewhat significant or very significant by neurophysiologist, thus these tests would have been performed.

Table 26 Comparison of referring physician's and neurophysiologist's opinion about ENMG test's significance, N=33. Referring physician’s evaluation of ENMG test’s significance for diagnosis decision-making.

		Neurophysiologist's prior evaluation of the significance of ENMG test			
		Insignificant	Slightly significant	Somewhat significant	Very significant
Referral physician's evaluation of test's significance for diagnosing	Insignificant	0 % (n=0)	0 % (n=0)	0 % (n=0)	0 % (n=0)
	Slightly significant	0% (n=0)	6.1% (n=2)	9.1% (n=3)	3% (n=1)
	Somewhat significant	0% (n=0)	0% (n=)	27.3% (n=9)	9.1% (n=3)
	Very significant	0% (n=0)	9.1% (n=3)	27.3% (n=9)	3% (n=1)
	Control test	0% (n=0)	3% (n=1)	0% (n=)	0% (n=)
	Don't know	0% (n=0)	0% (n=)	3% (n=1)	0% (n=)

Even though, referring physician’s opinion about ENMG test’s significance was in accordance with neurophysiologist’s opinion, the evaluation of quality of referral queries differed between the two groups. 50 % (n=16) of ENMG tests were evaluated by both parties as well-defined and clear. However, the remaining 50 % (n=16) of these referral queries were evaluated dissimilarly (Table 27 below); 9.4 % (n=3) of referral queries were evaluated as well-defined by referring physicians while neurophysiologists evaluated those to be poorly defined and unclear.

Table 27 Comparison of referring physician's and neurophysiologist's opinion about the quality of referral query

		Neurophysiologist's evaluation of referral query			
		Referral query missing	Poorly defined and unclear	Somewhat defined, but unclear and irrelevant	Relevant and well defined
Referring physician's evaluation of referral query	Referral query missing	0% (n=0)	0% (n=0)	0% (n=0)	0% (n=0)
	Poorly defined and unclear	0% (n=0)	0% (n=0)	3.1% (n=1)	0% (n=0)
	Well defined and clear	0% (n=0)	9.4% (n=3)	18.8% (n=6)	50% (n=16)
	Don't know	3.1% (n=1)	0% (n=0)	6.3% (n=2)	9.4% (n=3)

Referring physicians evaluated that for patient diagnosing 0 % of ENMG tests requested were insignificant and 14.3 % (n=10) were only slightly significant (Table 28 below). For treatment plan decision-making 2.9 % (n=2) were evaluated to be insignificant and 14.3 % (n=10) only slightly significant (Table 28). Altogether 11.4 % (8/70) were evaluated as insignificant or only slightly significant for both patient diagnosing and treatment plan decision-making. Referring physicians evaluated that very significant ENMG tests accounted for 45.7 % (n=32) for patient diagnosing and 32.9 % (n=23) for treatment plan decision-making (Table 28).

Table 28 Results of how referring physicians evaluated the significance of requested ENMG test

ENMG test result was significant for final diagnosis	
Insignificant	0% (n=0)
Slightly significant	14.3% (n=10)
Somewhat significant	37.1% (n=26)
Very significant	45.7% (n=32)
Control test	1.4% (n=1)
Don't know	1.4% (n=1)
ENMG test result was significant for treatment plan decision-making	
Insignificant	2.9% (n=2)
Slightly significant	14.3% (n=10)
Somewhat significant	50% (n=35)
Very significant	32.9% (n=23)
Don't know	0% (n=0)
Evaluation of referral query	
Referral query missing	2.9% (n=2)
Poorly defined and unclear	4.3% (n=3)
Well defined and clear	78.6% (n=55)
Don't know	14.3% (n=10)

To summarize the findings, referring physician's and neurophysiologist's opinion about the significance seemed to be mainly in line, but that the opinion about appropriate referral query differed. Referring physicians seemed to evaluate referral queries to be more appropriate than what neurophysiologists evaluated them to be. The number of inappropriate ENMG test referrals was according to neurophysiologists' evaluation 30.9 % (87/282) of all ENMG test referrals, while referring physicians evaluated that only 7.2 % (5/70) of their referrals to be poorly defined and unclear or missed referral query.

4.5 Controversial diagnosis groups

In this chapter the findings of interviews with experts are presented. There were three interviews made to specialists who were superior experts in their own field. Interviews included two selected patient diagnosis groups *morton neuroma* and *meralgia paresthetica*. These diagnosis groups were selected due to neurophysiologists' observation that ENMG tests in these patient groups are often difficult and unsuccessful to perform and the test result often remains equivocal. Furthermore, one interview was conducted to spine surgeon specialist in the attempt to find out how and when ENMG tests generally affect diagnosing and treatment plan decision-making in the field of spine orthopedics.

4.5.1 Morton neuroma

Interviews concerning *morton neuroma* patient diagnosis group revealed the following; The interviewee had the opinion that it is extremely important to have an ENMG test for decision-making before operation in a suspicion for *morton neuroma*. Practically, in the studied clinic, lesion release operation is not performed without ENMG test's proof of nerve lesion. Therefore, ENMG tests are not ordered if these test are not needed. For *morton neuroma* suspects, a clinical examination or a MRI is not enough to confirm the diagnosis or to decide on operation. Furthermore, ENMG tests are never requested without having seen the patient first, thus unnecessary and too general and unspecific ENMG requests are never made.

Of the 19 patient cases discussed, the interviewee evaluated that 100 % were either significant or very significant. The main reason why these ENMG tests were significant was that it either confirmed an operation or excluded an unnecessary *morton neuroma* operation.

4.5.2 Meralgia paresthetica

Unlike in *morton neuroma*, ENMG test conducted to *meralgia paresthetica* suspects were not always evaluated as significant or very significant. The interviewee stated that *meralgia paresthetica* diagnosis usually can be set by clinical examination. However, according to him, iatrogenic (nerve injury caused in operation) *meralgia paresthetica* suspects should always be examined with ENMG test.

Of the 17 *meralgia paresthetica* patient cases evaluated, 35 % were considered to be either significant or very significant and 47.1 % were evaluated as insignificant. The main reasons why ENMG test was not considered significant was that either ENMG test result was controversial to other diagnostic test results, pure clinical examination would have been enough to decide on diagnosis of *meralgia paresthetica*, or ENMG test report did not answer to the referral query. In addition to these reasons, an ENMG was considered unnecessary if patient's symptoms did not correlate with the finding or with anamnesis. The major reasons why ENMG test was significant for *meralgia paresthetica* suspects was that the test either specified diagnosis (33.3% of cases) or confirmed operation decision (33.3 % of cases). The two other reasons were that ENMG test excluded diagnosis (22.2 % of cases) or it provided new diagnosis alternative (11.1 % of cases).

4.5.3 ENMG tests' significance in spine patients candidates for spine surgery

According to the interview, occasionally diagnostic tests can be ordered inappropriately due to patient demands, meaning that patients often require a diagnostic test to explain their pain and symptoms. Often patients do not trust their physicians and the clinical examination, and they demand for more evidence through diagnostic tests.

In spine diagnostics, ENMG is not the primary diagnostic test and ENMG alone does not provide enough information. Therefore it is essential to relate ENMG test to other diagnostic tests such as x-ray test or MRI test. If ENMG was not available, the best alternative would be to do better clinical examination. According to the interviewee, ENMG tests are not requested without a proper indication, and only occasionally these tests are ordered without clinically investigating the patient before the ENMG request.

According to interviewee, specialists often consider zero result to be as important as other results from ENMG, and especially when zero result is helping with differential diagnosing. Specialists deem ENMG statements very important and they have to rely on these and to the terminology used in test result reports. Due to this if neurophysiologist do not answer clearly to the referral question, ENMG test may become useless. This leads to the fact that presenting a clear and well-defined referral query is also essential for referring physicians in order for them to get the most out of ENMG test.

The interviewee had the opinion that the primary care requests ENMG tests with less justifiable indications than the secondary care, which may be due to the poor understanding of the characteristics of ENMG test in the primary care. According to the interviewee, primary care may even request ENMG tests as primary diagnostic tests, although ENMG does not serve as such. It is crucial to first determine whether patient has peripheral or distal injury with MRI, and only after that send patient to ENMG test.

Of the discussed 20 patient cases, 20 % (n=5) were evaluated as significant or very significant. The main reason for ENMG test to be significant or very significant is when ENMG test confirms, specifies or excludes a diagnosis, and especially when ENMG test either confirms an operation-decision or helps to avoid unnecessary operation. Instead of requesting the 35 % (n=7) of patient cases discussed, which were evaluated insignificant ENMG tests, interviewee proposed an alternative to ENMG test. The alternative action compared to ENMG test was to do nothing. In most of these occasions ENMG did not provide new information or the result did not fit well to the information gathered from other diagnostic tests. Furthermore, a couple of these insignificant ENMG tests were considered useless, because the report did not answer to the referral question presented.

All interviewees stated similar characteristic of insignificant ENMG test which were:

1. ENMG test result is controversial to other diagnostic test results
2. ENMG test gave zero result, when it was not used for differential diagnostics
3. Test report does not answer to referral query
4. ENMG test is requested in order to win time either due to long queues to diagnostic tests and outpatient clinics or to put a patient to a test which has long waiting time giving the physician a time period to observe patient's symptoms whether they spontaneously disappear
5. ENMG test result does not correlate with clinical findings and symptoms
6. ENMG test is requested in order to find out what causes patient's symptoms or pain without any specific referral query
7. If thorough clinical examination would have confirmed diagnosis, then ENMG test could be considered as unnecessary.

Furthermore, the interviewed experts stated that the optimal referral query, that provides the most value to referring physician, is a question which can be answered with either 'yes' or 'no'. Neurophysiologists' own opinion about the patient case is valuable to referring physician. In addition, the referring physician usually appreciates possible suggestions given by the neurophysiologist concerning additional diagnostic tests.

5 Discussion

5.1 Summary of findings and conclusions

This chapter will present the conclusions based on the main findings of this research and relate them to findings of previous studies and literature. Managerial implications within the focal organization are presented. Related to managerial implications, the key challenges of behavioral change within healthcare are discussed. These key challenges present barriers, which have to be taken into account when implementing the selected approaches of demand management strategy. Lastly, this chapter will explore the limitations of this study with validity and reliability aspects.

To reiterate, this study has focused on exploring the factors affecting ENMG tests' significance and then on the strategies to reduce the number of insignificant ENMG tests. The topic was approached by conducting surveys to both neurophysiologists and referring physicians in order to clarify their opinion about significance of a ENMG test. These survey results were studied with statistical methods, supported by individual comments from interviews of three experts within the field. Furthermore, a literature review was conducted to gather information about the strategies used for reducing insignificant tests, to explore what causes the problem of increased demand, and how financial incentives have affected the behavior of requesting tests. Similarly, measures of effectiveness were explored, and selected to be used in this study. This research underlines propositions of how to reduce the number of insignificant ENMG tests. Statistical analysis and literature review all contributed to accomplish the objective of this study, which was to understand the factors affecting the significance of ENMG test and then describe the demand management strategies, which would be able to reduce the number of insignificant ENMG tests.

Firstly, characteristics of referral queries were found to be one of the main factors affecting ENMG tests' significance. The poorer the referral query, the less significant was the ENMG test result. Interestingly, the assumption that a poor referral query leads to a more extensive ENMG test could not be proved statistically. The effect of poor referral queries was identified in interviews as well, suggesting that ENMG test was insignificant to referring physicians too if the referral query was either missing, unfocused, or otherwise poorly defined. Interestingly, neurophysiologists and referring physicians evaluated referral queries differently, as even a half of the of cross-evaluated ENMG test referral queries were evaluated dissimilarly, and often so that the evaluation given by the neurophysiologist was poorer than that of the referring physician. In addition, neurophysiologist evaluated that poorly defined referral queries accounted 31 % of all ENMG test referrals, while the number was only 8 % for referring physicians. This finding suggests that increasing referring physicians' knowledge of well-defined referral query would be beneficial.

According to literature, demand management approaches to improve the quality of referral queries include redesigning the referral form, education, and constant feedback on referral query formulation to referring physicians (Fryer & Smellie 2013; Janssens 2010; Lang 2013; Lee et al. 2013). Furthermore, Spyridonis et al. (2013) presented that ‘*pain drawings*’ enable patients to describe their symptoms more objectively and precisely. This might make referral queries and a referral diagnoses more focused and clear. Currently, at HUS, ‘*pain drawings'* are drawn at the CN laboratory, and doing that already for the referral could be beneficial.

Secondly, neurophysiologists were found to be able to evaluate the significance of ENMG test based on test referral already. The analysis showed that the prior and final evaluations of significance of ENMG test were in line, meaning that the prior evaluation based on only test referral described the overall significance well enough. Besides, neurophysiologists’ prior evaluation was rather similar to referring physicians’ final evaluation of significance of ENMG tests. Hence, it would improve effectiveness of tests if neurophysiologists evaluated a greater number of ENMG test referrals and relied on their expertise while accepting and rejecting referrals. This suggest that neurophysiologists should continue their referral evaluation, and the practice of pre-approving referrals from certain referring quarters should be broadened.

In literature, evaluation and blocking of referrals is argued to be risky due to the possibility of blocking essential diagnostic tests (Janssens 2010), because laboratory personnel does not have enough information about the patient case. Nonetheless, Fryer & Smellie (2013) showed that assessing test requests decreased the share and cost of inappropriate tests. This method of reviewing test referrals should be adopted more widely within the field of ENMG testing, thanks to knowledgeable and capable neurophysiologists who are experts in their extremely specialized and advanced field of medicine.

Thirdly, other factors found lowering the significance of an ENMG test were test result’s expectedness and normality. The more surprising an ENMG test result, the more significant the test. A test did not have to provide surprising result in order it to be significant, but less significant tests more often provided unsurprising test results. Related to surprising ENMG findings, zero result ENMG tests were evaluated to be less significant than ENMG test that provided abnormal findings. However, interviews with experts revealed that zero results are essential for diagnosing a patient and deciding on treatment plan for a patient. That is to say, that a zero result is significant only if referring physician has requested the test for deciding between a couple of different diagnosis alternatives, called *differential diagnosing,* or if a zero result informed the referring physician to not operate on the patient.

In literature, opinions on normal test results and its significance vary (Shepherd 2010; Mondelli et al. 1998; Di Fabio et al. 2013; Cocito et al. 2006; Podnar 2005). Nonetheless, if diagnostic tests are requested too often, without clinical examination or interviewing a patient and ordered too many tests at a time, the probability of having normal findings increases, which may indicate inefficient resource utilization (Fuller 2005; Shepherd 2010; Podnar 2005). Furthermore, such an unsurprising zero result ENMG test could be avoided by doing a proper clinical examination to a patient before requesting ENMG. This phenomenon reflects to the formulation of a well-defined and relevant referral query, and according to Fuller (2005), a referring physician should not request an ENMG test, if he or she is not able to formulate a focused referral query. These findings are in line with our results.

Fourthly, the specialty of the referring quarter was found to affect the significance of an ENMG test. Especially primary care and physiatry outpatient clinics and rehabilitation centers requested less significant ENMG tests than other referring units. The possible reason to this is that primary care and physiatry outpatient clinics have different patient mix than other more specialized referring units. Patients in primary care and physiatry outpatient clinics have various health problems, and therefore physicians have a hard time identifying the cause of patients' symptoms. Our results support this assumption in that the referral diagnoses of physiatrists and primary care physicians more seldom agreed with the final finding. In this setting, formulating a perfectly focused referral diagnosis is more difficult than in other more specialized referring quarters. Possibly due to this, ENMG test referral queries from primary care and physiatry outpatient clinics are often unfocused, forcing neurophysiologist to conduct an extensive interview and clinical examination on a patient before even performing ENMG test. As a result, they also requested more ENMG tests with unsurprising, normal results.

Furthermore, unfocused ENMG tests may not provide valuable information, or at least not from neurophysiologist's perspective. It could be assumed, however, that such ENMG tests provide value to primary care or physiatry outpatient clinic physicians by still narrowing down diagnosis alternatives. Nevertheless, neurophysiologists are experts in conducting electrodiagnostic tests rather than taking a role as a referring physician. To optimize resource utilization, interviewing and clinically examining patients, as well as treatment pathway decisions, should be performed by the referring physician.

Besides, according to the survey for referring physicians, only 10 % of ENMG tests requested provided new diagnosis alternatives. This shows how few ENMG test actually help referring physician to diagnose a patient, if the physician has no referral diagnosis or idea of what might be the reason for patient's symptoms. In other words, ENMG test should not be requested if referring physician does not know how to focus and formulate referral query. It should not be used as a screening tool (Fuller 2005). According to our interviews, experts stated that especially primary care physicians are not all aware of the characteristics of ENMG testing, and they see it as primary diagnostic test. Our findings are in line with Fuller's (2005) argument that an ENMG test should be an extension to other diagnostic tests, not an independent primary test.

The greatest challenge with primary care is that the physicians there change constantly, because physicians are required to work a period of time in primary care as a part of their training in order to graduate. Due to this, educating them individually would be inefficient. However, education of primary care chief physicians might be a useful approach, because then the chief physician could educate the other primary care physicians. Furthermore, building an efficient and ongoing communication channel with primary care physicians would be valid approach to affect test requesting.

In addition to education and feedback, literature suggests financial incentives such as increased pricing for poorly defined referrals as a demand management strategy (Tierney et al. 1990; Seguin et al. 2002; Miyakis et al. 2006; Gama et al. 1992; Sood et al. 2007;Cummings et al. 1982). Overall, the results suggest that demand management efforts should be focused especially on physiatry outpatient clinics and primary care centers.

Although primary care physicians' referral diagnoses had lower agreement with the ENMG finding than secondary care referral diagnoses, primary care physicians were most often correct with the diagnoses of CTS and polyneuropathy. Mondelli et al. (2014) suggest that requests from primary care are approved only for the diagnosed CTS cases, and Mondelli et al. (1998) suggest that referrals from certain referring quarters should first be evaluated by a neurologist before the ENMG test. According to the results of the present study, CTS and polyneuropathy cases could be automatically approved from these referring quarters, while others should still be subject to neurophysiologists' approval. On the other hand, transferring and casting patients into other units' responsibility, i.e. in this case to neurology outpatient clinic, would only support the existence of current healthcare's 'silos'. Therefore, the approach suggested by Mondelli et al. (1998) would require further research.

Scholars argue that the value of ENMG test is highly related to neurophysiologists' skills, expertise and experience (Hellmann et al. 2005; Daube & Rubin 2009; Smith 2003; Chémali & Tsao 2005; So 2009). The results of the ordinal logistic regression model seemed to support this argument, as it indicates that the ENMG tests performed by interns were less significant than ENMG tests performed by specialists. However, the patient group examined by interns did not differ from patients examined by specialists. The underlying reason for this finding may not be the lower expertise of interns, but the different attitude of interns towards the evaluation of significance of ENMG test. In other words, interns seemed to be more willing to adopt the perspective that there in fact are insignificant ENMG tests, while specialists were very much against that thought. This same phenomenon of criticism towards research about quality was addressed by Mondelli et al. (1998), who presented that studies which judge competences of physician colleagues and concern quality assurance are often criticized by physicians. Similarly, this study invoked same kind of criticism both from neurophysiologists' and referring physicians' side. Therefore, the difference of significance of ENMG tests between interns and specialists is probably not solely depending on the expertise they have, but also the difference in attitudes towards evaluating one's own performance and work, and ability to question the current practices.

Finally, the percentage of *insignificant* ENMG tests performed was 6.7 % according to neurophysiologists' evaluation and 2.9 % according to referring physicians' evaluation. However, the number of *slightly significant* ENMG test was only 1.8 % according to neurophysiologists, and 14.3 % according to referring physicians. Altogether, the number of insignificant and only slightly significant ENMG tests were 10.3 %, 33 out of 319 all patient cases evaluated by either referring physicians (n = 70) or neurophysiologists (n = 249).

It is debatable which group is more capable of evaluating the actual significance and effectiveness of an ENMG test; referring physicians, who make treatment decisions and service the customer, or neurophysiologists, who are experts in the specific field of medicine. In the business world, firms tend to rely on the customer satisfaction and believe that 'the customer is always right' (Danneels 2003). Therefore one might consider referring physician's opinion concerning the significance of an ENMG test more representative. In addition, referring physicians are actually the ones that make the decisions to which the ENMG test is meant to provide supportive information. Therefore if referring physician has evaluated that ENMG did not provide effective information for the decision-making process, then it should be considered as the truth. On the other hand, one can argue that neurophysiologists are better at identifying the factors contributing to a successful and significant ENMG test. Neurophysiologists are experts in their field of medicine and they have performed numerous ENMG tests, thus they are capable of seeing when an ENMG test did not provide needed information and why.

Previous studies have found the percentage of insignificant ENMG tests to be between 17 - 28 % (Di Fabio et al. 2013; Mondelli et al. 1998; Mondelli et al. 2014; Cocito et al. 2006). However, the numbers are not quite comparable, since the criteria for insignificant ENMG test varies between studies and are not the same as those used in this research. In prior studies, the authors had defined tests to be insignificant if these gave normal result, test referral was poorly defined, test result did not affect treatment plan or diagnosis, or there was no correlation between working diagnosis, referral diagnosis and final ENMG finding. Furthermore, in their studies, ENMG tests' significance were not evaluated by neurophysiologist like in our study. However, our percentage of insignificant tests is within the same scale as the percentages from previous studies.

Agreement between referral diagnosis and final ENMG test finding varied between 40.5 – 71.9 % in previous studies (Di Fabio et al. 2013; Mondelli et al. 1998; Mondelli et al. 2014; Cocito et al. 2006; Podnar 2005), while in this study the agreement was only 38.5 %. This difference indicates that ENMG test referrals sent to Meilahti CN laboratory really need reviewing and improvement, in order to increase the agreement percentage, which is one of the spin offs of this study. Nonetheless, our results are in line with previous studies, as the agreement was lower in primary care physicians' referral diagnoses than in secondary care physicians' referral diagnoses.

5.2 Practical managerial implications of this study

Based on the findings of this research, Meilahti CN laboratory has implemented a set of actions to reduce the number of insignificant ENMG tests and improve effectiveness of ENMG testing. These managerial implications are presented in this chapter.

First, the finding that a poor referral queries result in less significant ENMG tests, was noticed in Meilahti CN laboratory already before conducting this research. Therefore Meilahti CN laboratory redesigned their test referral form by adding a compulsory question concerning the referral query and the level of lesion. However, this improvement has not yet seemed to increase the number of appropriate ENMG test referrals.

Second, the correlation between prior and final evaluation of ENMG test significance encouraged neurophysiologists to rely on their expertise while reviewing and accepting ENMG test referrals. Meilahti CN laboratory has been evaluating and accepting every test referral made by primary care physician during the past years. Nonetheless, as a spin off from this study, Meilahti CN laboratory has now encouraged neurophysiologists to evaluate referrals more critically and reject test referrals when needed. Furthermore, ENMG test referrals from neurology outpatient clinic has been taken for further evaluation for experimental reasons.

Because primary care physicians and physiatrists were found to order less significant ENMG tests, three different demand management strategies were adopted by Meilahti CN laboratory to reduce these insignificant tests. First, new, stricter instructions and limitations concerning ENMG test ordering were formulated and sent to primary care centers. For example, an approach of prohibiting ENMG test referrals by primary care physicians in other occasions than suspected carpal tunnel syndrome (CTS), has been proven effective in HUS Jorvi CN laboratory. Although strict instructions and limitations are being created for primary care, the strategy of HUS is to place itself closer to primary care, so that not too many patient end up in secondary care. Therefore, primary care physicians are instructed to seek help specialist help in form of consultations, instead of referring patients to secondary care. Furthermore, neurophysiologists in Meilahti established a helping phone line to advise primary care physicians in ordering of electrodiagnostic tests. Second, an educational meeting was set with primary care chief physicians. Third, new instructions for physiatrists concerning the ordering of ENMG tests are in progress and also an educational meeting with physiatrists has been decided.

Interestingly, Meilahti CN laboratory offered a test package for carpal tunnel syndrome (CTS) diagnosis for primary care physicians already before conducting this research. However, primary care physicians were reluctant to accept the package: they were afraid of underdiagnosing the patient by only ordering the carpal tunnel syndrome package. This may show that referring physicians want to avoid responsibility and prefer rely on neurophysiologists' professionalism. Especially in situations where a primary care physician is not sure of the diagnosis, they could certainly benefit from a helping phone line, which will probably prevent an insignificant ENMG tests to be requested.

Diagnosis groups of *meralgia paresthetica* and *morton neuroma* were explored in more detail to determine whether ENMG tests are useful in these patients. According to literature, diagnosis of *meralgia paresthetica* is commonly a result of clinical examination, although tumor and lumbar disc herniations must be excluded with diagnostic tests (Patijn et al. 2011). Our interviewee, an expert within this field, supported this proposition, and, save for the iatrogenic nerve injury leading to *meralgia paresthetica*, HUS could stop performing these ENMG tests.

Diagnosing *morton neuroma* with ENMG tests is difficult, as the test is difficult to perform to such patients. Thus, we explored the possibility of shifting towards the Swedish model, where *morton neuroma* is first tested with MRI test and then operated without performing ENMG test. However, according to our expert interviewee, the operation of *morton neuroma* without ENMG test is risky and may lead to unnecessary operations, which should be avoided. As a result, ENMG test was evaluated appropriate for *morton neuroma* patients, and that HUS does not operate patients without first having the ENMG test.

In the light of results that showed ENMG tests to be less significant when ordered by primary care physician or physiatrist, the differences between HUS CN laboratories become interesting. The percentage of ENMG tests requested by primary care physicians and physiatrists was much higher in Hyvinkää and Lohja CN laboratories than in Meilahti, Peijas or Jorvi CN laboratories. Furthermore, the ratio of ENMG test referrals from primary care physicians and physiatrists per population is higher in Hyvinkää and Lohja than in Meilahti. In addition, waiting time was also shorter in Hyvinkää and Lohja compared to Meilahti, Peijas and Jorvi CN laboratories. Due to the fact that in Hyvinkää and Lohja, a greater share of ENMG test referrals are made by physiatrists and primary care physicians, the potential of reducing less significant ENMG tests could be relatively larger than in Meilahti, Peijas or Jorvi CN laboratory. Besides, after the reduction of insignificant ENMG tests in Hyvinkää and Lohja, some patients could be even transferred from Meilahti to Hyvinkää and Lohja CN laboratories to shorten queues in Meilahti CN laboratory.

All in all, this study lead to numerous managerial implications within Meilahti CN laboratory, which are summarized in Table 29.

Table 29 Summary of research findings linked to managerial implications of Meilahti CN laboratory

Finding	Actions taken by HUS, *further implications*
Neurophysiologist can rely on his/her expertise on evaluating the significance of ENMG test based on test referral	• Neurophysiologists are encouraged to evaluate test referrals and reject them if needed • Test referrals from neurology outpatient clinic are being reviewed by neurophysiologist for experimental reason
Primary care and physiatry outpatient clinic order less significant ENMG test than other outpatient clinics	• New instructions with new limitations is written and sent to primary care centers • Neurophysiologists established a helping phone line for primary care physicians • An education session is set up with primary care chief physicians • New instructions are being written to physiatrists and given an presentation of instructions to them
Poor quality of referral query results in less significant ENMG test (usually to unsurprising zero results due to poor clinical examination and patient interview)	• New instructions are sent to both primary care and physiatry outpatient clinics • Education session with referring quarters are arranged constantly • The findings of this study are being communicated to neurophysiologists and emphasized the importance of answering to the referral query • *Pain drawings should be attached to referrals to improve referral quality and help neurophysiologist evaluate referrals*
At the moment, communication with referral physicians is minimal with no agile communication channels	• Neurophysiologists established a helping phone line for primary care physicians • Education session with referring quarters are arranged constantly • Communicating referring units about the regional distribution of demand
Meralgia paresthetica patients could be diagnosed in some situations with clinical examinations without ENMG	• New instructions are being written in HUS-level concerning the examination of *meralgia paresthetica* patients
According to literature, displaying test prices could reduce the number of inappropriate tests	• *If a neurophysiologists needs to perform an extensive clinical examination before the ENMG test due to a poor referral query, this could be charged separately from the referring quarter*
Primary care physicians are often correct concerning CTS and polyneuropathy referral diagnoses, but otherwise less often correct with referral diagnoses	• *CTS and polyneuropathy referrals could be automatically accepted, while subjecting other referrals from primary care physicians to prior evaluation*
Primary care physicians and physiatrists make up a larger share of ENMG test referrals in Hyvinkää and Lohja than in Peijas, Meilahti and Jorvi	• *There is a greater potential to omit insignificant tests in Hyvinkää and Lohja, possibly enabling transfer of tests to shorten queues in Meilahti*

5.3 Savings potential

Assuming that customer, here the referring physician, is the right person to evaluate which test provided value into decision-making process, the cost savings potential is calculated according to referring physician's opinion. Therefore, while ENMG tests that were evaluated insignificant or only slightly significant from both diagnosing and decision-making perspectives make up to 11.4 % of all ENMG tests performed, it results in roughly 460 insignificant or only slightly significant ENMG test per year in Meilahti CN laboratory, where approx. 4000 ENMG tests are performed annually.

One ENMG test takes about 30-60 minutes of neurophysiologist's time. Further, an insignificant test may take even longer time due to poorly defined and unfocused test referral. In addition, writing the ENMG test report takes additional 15-20 min. We can safely calculate that 460 tests account for 460 hours of neurophysiologist time plus the time for writing out the test result report to referring physicians being roughly 150 hours. In total, 610 hours of neurophysiologists' time is allocated to perform insignificant and only slightly significant ENMG tests annually in Meilahti CN laboratory. This amount of neurophysiologist time could be reallocated to perform very significant ENMG tests instead of insignificant. 610 hours of neurophysiologists' time account for 0.37 person-years (1 person-year is 220 working days). Of course, one has to deduct the time that it would take to evaluate ENMG test referrals and decide, which should be prioritized in order to perform fewer insignificant ENMG tests. The total number of ENMG tests performed a year in the HUS area was 8550 in year 2014, which means 980 (11.4 %) insignificant or only slightly significant tests, accounting to approximately 980 hours only to perform the tests, plus the time to write the report accounting to 330 hours. In total, roughly 1300 hours of neurophysiologist time is allocated annually in HUS CN laboratories to insignificant ENMG tests, transferring to 0.79 person-years, which could be reallocated more effectively.

In monetary terms, insignificant and only slightly significant ENMG tests account for roughly 154 000 € per year in Meilahti CN laboratory, while one ENMG test costs 337 €, and assuming that the pricing of an ENMG test presents all costs caused by performing one ENMG test. Cost savings in total for all CN laboratories within HUS area altogether it would be roughly 329 000 € per year, if the 11.4 % of all ENMG tests accurately presents the share of insignificant ENMG test in other CN laboratories as well.

However, even though the aim of this research is to reduce the number of insignificant tests, the greater goal is to reallocate CN laboratory resources so that they are able to provide more value to customers i.e. referring physicians and patients; allocating neurophysiologists time to very significant ENMG tests rather than insignificant and only slightly significant tests. According to the results of this research, a better clinical examination done by referring physician and then formulating a more focused, relevant and well-defined referral query would save a lot of resources. Above all, a good clinical examination might even hinder the need for an ENMG or some other test. On the other hand, this raises the question, what to do with the patient instead of the ENMG test?

5.4 Challenges in deploying changes within current healthcare system

The managerial implications presented in previous chapter embody actions that are meant to change the behavior of referring physicians and replace ENMG tests with other actions. However, replacing actions must be explored and achieving a change within current healthcare system is challenging, which are discussed in this chapter based on the literature review and findings of this study.

As stated in previous chapter, omitting assumedly insignificant ENMG tests only creates savings if they are replaced with more cost-effective alternatives. If a referring physician does not know what causes patient's symptoms even after careful clinical examination and he is not able to formulate a good referral query, then he or she might have to refer patient to a specialist of neurology i.e. neurologist. This of course would increase queues for neurology outpatient clinic, but it might also improve patient outcome. This is improved as patients get correct treatment within shorter period of time than what they would have gotten out of having first an insignificant ENMG test and only then referred to neurology outpatient clinic. However, according to interviews, a valid alternative to requesting an insignificant ENMG test was doing nothing to the patient. Therefore, although suggested by Mondelli et al. (1998), the referral to neurology outpatient clinic in such situations might not be recommended either.

Generally, referring patients into other 'silos' of the healthcare system is not the answer to long queues. However, currently the Finnish healthcare system functions so that patients are constantly circulating between different healthcare units. This is especially true in patients with multiple difficult-to-treat and difficult-to-diagnose health problems. However, the demand stays the same or even increases if we continue to 'cast patients' to other silo's responsibility. A better solution would be to have functioning communication channels between healthcare sectors and 'silos', and to have more integrated healthcare system, which would pursue towards common goal.

Generally, having long waiting times to diagnostic tests creates a vicious cycle. According to interviews, referring physicians refer patients to multiple diagnostic tests at the same time in order to save time in waiting and to begin patient's treatment as early as possible. The referring physician first evaluates the reports of all these tests only before the next patient visit. This obviously results in unnecessary tests, when one of the tests ordered already confirms the diagnosis, making the other tests redundant. This not only delays the actual treatment, but also increases waiting times for other patients. Shorter waiting times would enable referring physicians to rely on requesting only the most appropriate test, one at a time, evaluating the results before ordering more tests.

However, making behavioral changes in physicians could prove challenging, because these changes should optimally be orderer centric, not producer centric. Part of the culture in Finnish healthcare industry, most physicians are inherently inclined to always think what's best for the patient at hand, regardless of the costs. Furthermore, the new trend of consumers extensively measuring and tracking their health and bodily functions, called *quantified self* (Swan 2012), has increased demand for diagnostic tests and shifted attitudes towards demanding knowledge, instead of relying on the professionalism of a physician. A good example of this is the rising market for specialized diagnostic tests provided by private firms, such as Cityterveys Oy operating in 5 locations in Finland, advertising tests to anyone willing to pay. These kinds of intrinsic motivators for physicians' behavior need actions that create orderer centric motivation, such as financial motives, instead of producer centric changes.

From the business strategy perspective, it is unusual that demand is sought to be decreased, not increased. Although perceptions are dissimilar in demand management literature, the business value of information is still important also in this situation. Value of information could be evaluated by stating a question of: "How much money would a referring physician be willing to pay for the test, if he will be rewarded a greater sum for a test that turned out significant?" In other words, referring physicians would need to have their own individual budgets instead of budgets for the whole referring unit or municipality. Referring units, e.g. a primary care unit would have internal budget allocation between their physicians and physician's budget would be affected if they order tests or buy healthcare services for their patients. In that way they would become well aware of the costs and not have financial information asymmetry between the diagnostic test performing unit. However, personal budgeting would increase administrative costs and could create dilemmas when e.g. the budget runs out (Janssens 2010). Furthermore, there are other unwanted outcomes of this kind of financial incentive creation such as seen in NHS's QOF-system, where physicians began to select patients according to their monetary benefits (Doran et al. 2011; Roland 2004; Mangin & Toop 2007). However, some state that if incentives are not changed, nothing will change (Bernardy et al. 2009).

The increased demand is difficult to control due to strong incentive asymmetry, where different parties have different goals. Referring physicians want to solve patient's case as soon as possible, with as many diagnostic tests as it requires and with high quality not depending on the cost of care. On the other hand, the diagnostic test provider wants to conduct tests to only patients with actual need for the test and to whom it brings the most value, and they want to avoid unnecessary costs and long waiting times.

Reducing the number of inappropriate tests would create cost savings, allocate resources more effectively, shorten waiting times, improve referring physicians diagnosing and treatment plan decision-making, and thus increase quality of care. If patients were diagnosed and treated earlier, it would improve the overall patient outcome. Therefore, reducing insignificant diagnostic tests would improve healthcare system in many sectors rather than only in diagnostic testing. These demand management approaches and factors affecting insignificant ENMG testing can be replicated to other fields of diagnostic testing, e.g. radiology and other medical imaging. In addition, some of the findings could apply in other fields of medicine outside diagnostic testing, for example regarding financial incentives and alternative policies in referring patients from primary care to secondary care.

5.5 Reliability, validity and limitations of the study

This chapter presents the reliability and validity of this study and discusses the possible limitations.

Reliability is measured through whether a research is possible to be replicated and that this replicated study provides similar results (Flick 2009). Research methods used in this study are presented in detail, and justified reasons for the selection of the specific methods used are presented. This makes the present study possible to replicate in other settings. In terms of data accuracy, the reliability of this study is slightly limited due to the rather limited samples used. While all findings presented were shown to be statistically significant with high confidence levels, the strengths of those effects are subject to measurement errors. This research should be replicated with larger samples before making decisions based on exact numeric results of this study, especially in other settings outside HUS and Meilahti CN laboratory.

Construct validity describes whether the study measures the studied phenomenon with logical and reasonable metrics and methods (Flick 2009). This study conducted a wide literature review from which suitable metric to measure the significance and effectiveness of a diagnostic test was selected. The selection was further confirmed through expert interviews. While choosing the right construction for measuring significance and effectiveness healthcare is largely a value based problem, we have considered the problem from multiple perspectives, using multiple methods, and chosen ones that suit the focal organization. Thus, we argue that construct validity is high in this study.

External validity measures the extent to which a research's findings can be generalized (Flick 2009). Because this study focused on exploring the effectiveness electrodiagnostic tests and especially ENMG tests, the findings of this study can at widest to be generalized to radiological tests and with some carefulness to other diagnostic testing. However, the findings of the literature review partly apply for a wider context, and our findings provide sound suggestions for testing in other contexts. In order to increase the external validity and to generalize the findings of this study to a wider context, additional similar studies are needed in other sectors of healthcare.

The ability of our methods to measure the concept of significance of ENMG testing can be debated. The surveys conducted to both neurophysiologists and referring physicians may have resulted in biased answers, because healthcare professionals have been found to act in a protective manner when their work performance is observed (Mondelli et al. 1998). Hence, when introducing these surveys, some criticism was presented toward this study setup, because the aim of the study was to ask the physicians to evaluate the significance of tests they have either requested or performed. One improvement would have been to get a second opinion from another physician concerning the same patient cases. However, this study now presents physicians' opinions, who are locally involved in the process of requesting and performing these ENMG tests. Because the opinions are gathered from these physicians, the number of insignificant cases found is likely based on conservative evaluations, and thus the results are more likely to be also accepted by the physicians involved in the process, making the results of this study more useful. Besides, decision-making concerning patient's diagnosis or treatment plan is done according to physician's own opinions about the results of diagnostic tests. Therefore, our results present this decision-making process and effectiveness of these tests very well.

A second limitation was the small number of data from referring quarter survey made to primary care physicians. Therefore, the results from that survey may not be very representative. This was a challenge not easily overcome, since referring physician turnover is high within a primary care unit, making them hard to reach. The survey made to neurophysiologist included ENMG tests requested by primary care physicians, however, thus this subproject managed to gather at least some opinion about primary care requests. The survey was made to primary care physicians too, but unfortunately they request ENMG tests only occasionally during one month and such a primary care unit has many physicians, thus gathering their answers was not possible. Furthermore, we gathered few patient cases which were evaluated by both neurophysiologists and referring physicians, therefore no statistical tests could be employed to verify the observations.

Thirdly, the lack of purely quantitative data analysis of ENMG test usage and patients' treatment pathways could be considered one limitation of this study. However, such data was unfortunately not available since it is not recorded in the first place. This issue is at present under development, and treatment pathway data recording will be implemented in a future patient information system in HUS. This research only gathered data concerning physicians' opinions, which are subjective. The answers of neurophysiologists' survey were checked with chi-square test not to differ between neurophysiologists, thus the opinion about the significance of ENMG test can be considered as consistent and reliable.

Lastly, the final coding of survey answers is partly dependent on the coder, thus categorization of answers within this study is one perspective to examine the findings. However, the used coding was verified many times before calculating the final results, and findings gathered from quantitative and qualitative content analysis can be considered as reliable. Nonetheless, due to fact that significance was measured in this study and was the focus of analysis, one has to take into account that the interpretation of the word 'significant' may vary between physicians who answered to the survey, even though the survey included a definition of significance.

Overall, despite some limitations of this study, our findings are mainly analogous to the results of similar previous studies, giving support to our results.

6 Summary

6.1 Answers to research questions

This chapter summarizes the findings of this study and provides the answers to research questions presented in chapter 1.2. Furthermore, the future research areas and topics evoked by this study are discussed.

The main aim of this study was to identify strategies to reduce the number of insignificant diagnostic tests in healthcare. First, we identified characteristic of an insignificant diagnostic test, and then explored demand management strategies through a literature review to reduce the number of insignificant ENMG tests. Lastly, the application of those methods were explored in the case of HUS and its Meilahti CN laboratory. Next, each of the questions and the answers to them are presented individually.

What strategies could be used to reduce the number of insignificant diagnostic tests in healthcare?

To answer this question, first, a literature review was conducted in order to explore the demand management strategies used in previous studies to reduce the usage of insignificant diagnostic tests. Second, the characteristics of an insignificant diagnostic test were researched from prior studies and a survey conducted to neurophysiologists, to find the most valid demand management strategies appropriate for this setting. Lastly, the reasons for increased demand of diagnostic tests were searched from literature and the performance data of Meilahti CN unit was analyzed.

The main demand management strategies presented in literature were meant for diagnostic test service providers to be implemented into their operations. These approaches included educating test requesting physicians, giving feedback, referral evaluation, redesigning the test request form and establishing constant communication with test orderers (Janssens 2010; Fryer & Smellie 2013; Mondelli et al. 1998). Demand management strategies requiring political decision-making and a mutual agreement with service requestor were blocking diagnostic tests from selected referring quarters (Mondelli et al. 1998; Fuller 2005) and creating financial incentives for referring physicians (Fryer & Smellie 2013; Mannion 2014; Doran et al. 2011; Bernardy et al. 2009; Marshall & Smith 2003).

However, studies have shown that education and feedback giving is expensive (Schroeder et al. 1984) requiring a lot of resources to produce feedback and have educational sessions (Janssens 2010), and the savings achieved have been marginal and consumed by the costs of education and giving feedback (Eisenberg 2002; Schroeder et al. 1984). Nonetheless, providing information, feedback and communication channels is not as expensive nowadays as it has been in the past, due to advanced IT-systems.

Two of the main factors found to predict the significance of an ENMG test were the quality of test referrals and the referring quarter. Poorly defined referrals lead into less significant tests, and primary care physicians and physiatrists request less significant ENMG tests than other categories of specialists.

Therefore, more critical prior evaluation of test referrals, attaching ‘pain drawings’ to referrals already, and educating referring physicians about the characteristics of good ENMG test referrals are suggested in order reduce the number of insignificant ENMG tests. These measures should be focused on primary care physicians and physiatrists. Mondelli et al. (1998) suggested that ENMG test referrals from primary care should not be accepted in any other than CTS suspect case, which has been tested successfully in HUS Jorvi CN unit. Thus, the significance of ENMG tests requested by primary care physicians could be improved by education, reducing requesting possibilities (e.g. limiting primary care physicians to only requesting ENMG tests for CTS patients), or even charging an extra price for additional clinical examinations resulting from poor referral queries.

According to literature, while changes can be achieved with aforementioned demand management strategies, the most effective strategy is to create incentives for diagnostic test requestors (Bernardy et al. 2009; Kaplan & Porter 2011). Creating motivation is more effective in creating behavioral changes than forcing them with policies. For example, financial incentives would create motivation already at the requestor side. However, creating such monetary incentives is risky and might affect overall quality of care. For example, a UK study showed that creating economic incentives for physicians encouraged them to select patients who would ensure them greater income (Fleetcroft et al. 2012). Furthermore, creation of such incentives would require political impact and a lot of planning before implementation. In the end, these financial incentives have to be considered, because no significant change is likely to occur if the change depends only on the actions implemented by service provider (Bernardy et al. 2009).

What are the characteristics of an insignificant ENMG test?

- *What are the main factors contributing ENMG tests’ insignificance?*
- *Does the opinion about the ENMG tests’ significance or the quality of test referral query vary between referring physicians and neurophysiologists?*

To answer this question, a survey was conducted to neurophysiologists and perspectives of various stakeholders on the characteristic of an insignificant ENMG test were gathered through a literature review. First, the prior evaluation of ENMG test significance, made based on only ENMG test referral, showed to predict the final significance of an ENMG test well. This indicates that neurophysiologists are capable of evaluating the significance of an ENMG test according to a test referral.

Second, according to analysis, the poorer the referral query, the less significant the test was performed. The survey for referring physicians showed that only 10 % of ENMG tests requested provided new diagnosis alternative, indicating that requests made in hope of receiving new suggestions were rarely significant. An ENMG test should not be explorative, but instead, sufficient information should be gathered before the ENMG test (Di Fabio et al. 2013). Moreover, an answer to a focused and well-defined referral query provides more usable, practical information and thus more significant tests, a finding in line with prior studies (e.g. Fuller 2005; Di Fabio et al. 2013).

Third, according to neurophysiologists, normal and expected ENMG findings were less significant in average. Requesting an ENMG without a well-informed referral query more likely results in expected and normal ENMG findings. Literature presented varying opinions about the significance of normal or zero results (Shepherd 2010; Mondelli et al. 1998; Di Fabio et al. 2013; Cocito et al. 2006; Podnar 2005), but according to our interviews, a normal test result can be significant while used for excluding a diagnosis or for differential diagnosing. Furthermore, expected test results are significant when these tests provide information about post-operative symptoms. In conclusion, normal results are not an indication of insignificance.

Fourthly, primary care physicians and physiatrists requested less significant tests compared to other physicians. Referral diagnoses of primary care physicians and physiatrists agreed with ENMG finding less frequently than those of other physicians. They also requested more normal and unsurprising ENMG tests. This may be due to the challenging patient mix these quarters have, patients who have a wde scale of health problems and thus are difficult to treat, or excessive use of ENMG tests as explorative diagnosing (Di Fabio et al. 2013).

To answer to the second sub-question, the answers from neurophysiologists and referring physicians were cross-checked. The opinions of neurophysiologists and referring physicians about the significance of an ENMG test were mostly similar. Especially neurophysiologists' prior evaluation and referring physicians' final evaluation of significance were mainly in line. This indicates that neurophysiologists could rely on their expertise while evaluating the significance of ENMG tests according to test referrals. However, due to the small sample size, this finding could not be proved statistically.

Interestingly, test requestor and provider evaluated the quality of ENMG test referrals differently. Around 30 % of all ENMG test referrals were considered poor by neurophysiologists, while referring physicians evaluated them to account for only 7 %. Neurophysiologists were more critical of the referral queries that referrers rated as good, indicating the need of education for referring physicians.

> *What number of ENMG test requests and ENMG tests are insignificant in Meilahti CN laboratory and do referring units differ in this?*

As mentioned above, according to neurophysiologists, 30% of referral queries were poor. This means that the Meilahti CN laboratory receives approx. 1 200 poorly and inappropriately formulated ENMG test referrals annually.

About 11 % of all ENMG tests were evaluated to be either insignificant or only slightly significant by referring physicians, which accounts for 460 insignificant or slightly significant ENMG tests performed yearly in Meilahti CN laboratory. It was calculated that 460 tests account for in total 610 hours or 0.37 person-years of neurophysiologists' time that could be reallocated to perform more significant ENMG tests, reducing waiting times. For the HUS area in total, assuming similar share of insignificant tests, 11 % would translate into 980 tests, or 0.79 person-years that could be reallocated more effectively.

In monetary terms, calculated through the full costs of an ENMG study, insignificant and only slightly significant ENMG tests account for 154 000 € per year in Meilahti CN laboratory. For the whole HUS area, assuming a similar share of insignificant tests, the figure would be 329 000 € annually.

6.2 Future research

Although some previous studies about insignificant diagnostic tests exist, research concerning the utilization of such findings is minimal. The issue of effectiveness in healthcare becomes more and more important in future, and it is likely that various players within healthcare will begin to use the results of these kind of studies. In order to deepen the knowledge of insignificant tests and to clarify how to reduce these tests, this study found interesting future research possibilities.

Prior studies on effectiveness in healthcare services have focused on the characteristics of ineffective services. However, there are only a few studies which have evaluated the effect of remedies on service effectiveness and success of the adopted demand management strategies. Therefore, more 'before-and-after' studies of effectiveness are needed to find the most effective demand management strategies for healthcare. For the current topic, this would mean implementing the managerial implications suggested by this study, following closely the development of waiting times and number of tests requested, and finally conducting similar surveys concerning the significance of ENMG tests. In healthcare, empirical tests are challenging for ethical reasons, thus it is rarely feasible to test the effectiveness of single strategies in isolation.

The interesting finding of primary care physicians requesting less significant tests, requires further research on primary care physicians' requesting behavior. This study did not manage to gather many survey answers from primary care physicians. Therefore, a study focusing on this group of referring physicians is needed. It would be important to have their stronger opinion about the significance of ENMG tests. Previous studies have explored the differences between general practitioners and specialists (Mondelli et al. 1998), but these differences are usually based solely on neurophysiologists' opinions.

Comparison of evaluations is essential in order to understand how the valuable information is created and transferred to customer. This study managed to gather only 33 cross-evaluated patient cases. Thus a study with a larger number of comparable patient cases is needed to statistically verify the findings of this study. Such studies, comparing referring physicians' and test provider physicians' opinions about service effectiveness and significance, are missing in the current literature.

This study gathered data with surveys and questions presented to the actual decision-makers. However, there may be a possible bias of subjectivity and difficulty to critically evaluate one its own actions. Therefore, a similar study would be needed including third party actors who would give their second opinion. This would minimize the possible bias as the data could be cross-checked.

Finally, this study brought an interesting study topic concerning the role of financial incentives in Finnish healthcare system into the spotlight. Due to the fact that the current healthcare system needs a change in course, a further study of appropriate economic incentives is required in the future and in welfare countries like Finland. Furthermore, a main challenge remains that measuring performance and effectiveness is challenging within diagnostic tests, and requires also value-based decision-making. Performance measurement is important in motivating change, thus some metrics need to be decided on. These kind of studies would be needed to explore the ways to achieve behavioral change, because it will require a lot more than only healthcare service provider's actions. Namely, also healthcare service orderers have to change their behavior. Numerous such studies have been performed in United States and United Kingdom, but not in Finland, which still presents a fundamentally different healthcare system.

List of references

Allan, G.M. & Lexchin, J., 2008. Physician awareness of diagnostic and nondrug therapeutic costs: a systematic review. *International journal of technology assessment in health care*, 24, pp.158–165.

Anderson, G.F. et al., 2006. Health care spending and use of information technology in OECD countries. *Health Affairs*, 25(3), pp.819–831.

Bailey, J., Jennings, a & Parapia, L., 2005. Change of pathology request forms can reduce unwanted requests and tests. *Journal of clinical pathology*, 58(8), pp.853–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1770878&tool=pmcentrez&rendertype=abstract [Accessed October 20, 2014].

Bairstow, P.J. et al., 2006. Diagnostic imaging pathways : development , dissemination , implementation , and evaluation. , 18(1), pp.51–57.

Baxter, P. & Jack, S., 2008. Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *Qual Rep*, 13(4), pp.544–559.

Bender, R. & Grouven, U., 1997. Ordinal logistic regression in medical research. *Journal of the Royal College of Physicians of London*, 31(5), pp.546–551.

Bernardy, M. et al., 2009. Strategies for Managing Imaging Utilization. *JACR Journal of the American College of Radiology*, 6(12), pp.844–850. Available at: http://dx.doi.org/10.1016/j.jacr.2009.08.003.

Bertakis, K.D. & Azari, R., 2011. Patient-centered care is associated with decreased health care utilization. *Journal of the American Board of Family Medicine : JABFM*, 24, pp.229–239.

Blackmore, C.C., Mecklenburg, R.S. & Kaplan, G.S., 2011. Effectiveness of clinical decision support in controlling inappropriate imaging. *JACR Journal of the American College of Radiology*, 8(1), pp.19–25. Available at: http://dx.doi.org/10.1016/j.jacr.2010.07.009.

Bossuyt, P.M.M. et al., 2012. Beyond diagnostic accuracy: The clinical utility of diagnostic tests. *Clinical Chemistry*, 58(12), pp.1636–1643.

Bossuyt, P.M.M. & McCaffery, K., 2009. Additional patient outcomes and pathways in evaluations of testing. *Medical decision making : an international journal of the Society for Medical Decision Making*, 29, pp.E30–E38.

Bunting, P.S. & Van Walraven, C., 2004. Effect of a controlled feedback intervention on laboratory test ordering by community physicians. *Clinical chemistry*, 50(2), pp.321–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14684622 [Accessed January 12, 2015].

Campbell, S.M. & Lester, H., 2008. The Experience of Pay for Performance in English Family Practice : A Qualitative Study ABSTRACT. *Annals Of Family Medicine*, pp.228–234.

Chalkidou, K. et al., 2009. Comparative effectiveness research and evidence-based health policy: experience from four countries. *The Milbank quarterly*, 87(2), pp.339–367.

Chambers, J.D., Neumann, P.J. & Buxton, M.J., Does Medicare have an implicit cost-effectiveness threshold? *Medical decision making : an international journal of the Society for Medical Decision Making*.

Chémali, K.R. & Tsao, B., 2005. Electrodiagnostic testing of nerves and muscles: When, why, and how to order. *Cleveland Clinic Journal of Medicine*, 72(1), pp.37–48.

Cocito, D. et al., 2006. A further critical evaluation of requests for electrodiagnostic examinations. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 26(6), pp.419–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16601934 [Accessed September 18, 2014].

Creswell, J.W., 2014. The Selection of a Research Approach. In *Research Design*. pp. 3–23. Available at: http://www.sagepub.com/upm-data/55588_Chapter_1_Sample_Creswell_Research_Design_4e.pdf.

Cummings, K.M. et al., 1982. The effects of price information on physicians' test-ordering behavior. Ordering of diagnostic tests. *Medical care*, 20(3), pp.293–301.

Danneels, E., 2003. Tight-loose coupling with customers: The enactment of customer orientation. *Strategic Management Journal*, 24(6), pp.559–576.

Daube, J.R. & Rubin, D.I., 2009. Needle electromyography. *Muscle & nerve*, 39(2), pp.244–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19145648 [Accessed September 3, 2014].

Doran, T. et al., 2011. Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. *BMJ (Clinical research ed.)*, 342, p.d3590.

Driskell, O.J. et al., 2012. Inappropriate requesting of glycated hemoglobin (Hb A1c) is widespread: Assessment of prevalence, impact of national guidance, and practice-To-practice variability. *Clinical Chemistry*, 58, pp.906–915.

Drummond, M., Griffin, A. & Tarricone, R., 2009. Economic evaluation for devices and drugs--same or different? *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 12(4), pp.402–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19138306.

Eichler, H.G. et al., 2004. Use of cost-effectiveness analysis in health-care resource allocation decision-making: How are cost-effectiveness thresholds expected to emerge? *Value in Health*, 7(5), pp.518–528.

Eisenberg, J.M., 2002. Physician utilization: the state of research about physicians' practice patterns. *Medical care*, 40(11), pp.1016–35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12409848 [Accessed January 9, 2015].

Di Fabio, R. et al., 2013. Requests for electromyography in Rome: a critical evaluation. *Functional neurology*, 28(4), pp.281–4. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3951256&tool=pmcentrez&rendertype=abstract.

Felix, E.R. et al., 2010. Utility of quantitative computerized pain drawings in a sample of spinal stenosis patients. *Pain Medicine*, 11, pp.382–389.

Ferrante di Ruffano, L. et al., 2012. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *Bmj*, 344(February), pp.e686–e686.

Ferrante di Ruffano, L. et al., 2009. *How do tests impact on patient health? An explanatory framework*,

Fleetcroft, R. et al., 2012. Incentive payments are not related to expected health gain in the pay for performance scheme for UK primary care: cross-sectional analysis. *BMC Health Services Research*, 12(1), p.94. Available at: http://www.biomedcentral.com/1472-6963/12/94.

Flick, U., 2009. *An Introduction to Qualitative Research* 4th ed., SAGE Publications Ltd.

Flodgren, G. et al., 2011. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane database of systematic reviews (Online)*, (7), p.CD009255.

Fryback, D. & Thornbury, J., 1991. The efficacy of imaging. *Medical Decision Making*, 11, pp.88–94.

Fryer, A. a & Hanna, F.W., 2009. Managing demand for pathology tests: financial imperative or duty of care? *Annals of clinical biochemistry*, 46, pp.435–437.

Fryer, A. a & Smellie, W.S. a, 2013. Managing demand for laboratory tests: a laboratory toolkit. *Journal of clinical pathology*, 66(1), pp.62–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23015659 [Accessed September 18, 2014].

Fuglsang-Frederiksen, a et al., 1995. Variation in performance of the EMG examination at six European laboratories. *Electroencephalography and clinical neurophysiology*, 97(6), pp.444–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8536597.

Fuller, G., 2005. How to get the most out of nerve conduction studies and electromyography. *Journal of neurology, neurosurgery, and psychiatry*, 76 Suppl 2, pp.ii41–46. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1765696&tool=pmcentrez&rendertype=abstract [Accessed November 6, 2014].

Gama, R. et al., 1992. Modifying the request behaviour of clinicians. *Journal of clinical pathology*, 45(3), pp.248–249.

Gazelle, G.S. et al., 2011. A framework for assessing the value of diagnostic imaging in the era of comparative effectiveness research. *Radiology*, 261(3), pp.692–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22095993.

Goehler, A. & Gazelle, G.S., 2014. Examining the Use of Comparative and Cost-Effectiveness Analyses in Radiology.

Guthrie, B., Mclean, G. & Sutton, M., 2006. B Guthrie , G McLean and M Sutton Workload and reward in the Quality and Outcomes Framework of the 2004 general practice contract. *General Practice*, (September), pp.836–841.

Heikkilä, H.T. & Oja, S., 2015. ENMG-tutkimuksen asema hermoperäisen oireen selvittämisessä. *Neurotutka Oy*. Available at: http://www.neurotutka.fi/Neurotutka/ENMG-artikkeli.html [Accessed May 27, 2015].

Hellmann, M. et al., 2005. Diagnostic value of electromyography in children and adolescents. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 22(1), pp.43–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15689712.

Helsingin ja Uudenmaan Sairaanhoitopiiri, 2012. *Edelläkävijä - vaikuttavaa hoitoa potilaan parhaaksi. Strategia 2012-2016*,

Hoey, J. et al., 1982. Physician Sensitivity to the price of diagnostic tests: A U.S.-Canadian analysis. *Medical care*, 20(3), pp.302–307.

Janssens, P.M.W., 2010. Managing the demand for laboratory testing: options and opportunities. *Clinica chimica acta; international journal of clinical chemistry*, 411(21-22), pp.1596–602. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20659442 [Accessed October 13, 2014].

Johannesson, M. & Jönsson, B., 1991. Economic evaluation in health care: is there a role for cost-benefit analysis? *Health policy (Amsterdam, Netherlands)*, 17, pp.1–23.

Kaplan, R.S. & Porter, M.E., 2011. The Big Idea - How to Solve The Cost Crisis In Healt h Care. *Harvard Business Review*, (September), pp.46–64.

Karadag, Y. et al., 2014. Referral Diagnosis Versus Electroneurophysiological Findings-Three Years Experience from a Tertiary Hospital. , 11(4), pp.244–247.

Kernick, D.P., 1998. Economic evaluation in health: a thumb nail sketch. *Bmj*, 316(7145), pp.1663–1665. Available at: http://www.bmj.com/cgi/doi/10.1136/bmj.316.7145.1663.

Komur, M., Okuyaz, C. & Makharoblidze, K., 2014. Consistency between referral diagnosis and post-ENMG diagnosis in children. *JPMA. The Journal of the Pakistan Medical Association*, 64(2), pp.179–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24640809.

Kothari, M.J. et al., 1995. Electromyography: do the diagnostic ends justify the means? *Archives of physical medicine and rehabilitation*, 76(10), pp.947–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7487436.

Krippendorff, K., 2003. Content Analysis: An Introduction to Its Methodology. *Content Analysis: An Introduction to Its Methodology*, p.440. Available at: http://www.uk.sagepub.com/textbooks/Book234903.

Lacey, A. & Luff, D., 2007. Qualitative data analysis. *The NIHR RDS for the East Midlands / Yorkshire & the Humber*, pp.13–14.

Lang, T., 2013. Laboratory demand management of repetitive testing--time for harmonisation and an evidenced based approach. *Clinical chemistry and laboratory medicine : CCLM / FESCC*, 51(6), pp.1139–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23420284 [Accessed October 13, 2014].

Lee, A.C.K. et al., 2013. Demand management: misguided solutions? *The British journal of general practice : the journal of the Royal College of General Practitioners*, 63(613), pp.e589–91. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3722836&tool=pmcentrez&rendertype=abstract [Accessed October 13, 2014].

Lee, D.W., Neumann, P.J. & Rizzo, J. a, 2010. Understanding the medical and nonmedical value of diagnostic testing. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 13(2), pp.310–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19744295.

Lijmer, J.G., Leeflang, M. & Bossuyt, P.M.M., 2009. Proposals for a phased evaluation of medical tests. *Medical decision making : an international journal of the Society for Medical Decision Making*, 29, pp.E13–E21.

Loop, J.W. & Lusted, L.B., 1978. American Collage of Radiology Diganostic Efficacy Studies. *American Roentgen Ray Society*.

Loop, J.W. & Lusted, L.E., 1978. American College of Radiology Diagnostic Efficacy Studies. *AJR. American journal of roentgenology*.

Lysdahl, K.B. & Hofmann, B.M., 2009. What causes increasing and unnecessary use of radiological investigations? A survey of radiologists' perceptions. *BMC health services research*, 9, p.155.

Malterud, K., 2001. Qualitative research : standards , challenges , and guidelines. *The Lancet*, 358(panel 2), pp.483–488.

Mangin, D. & Toop, L., 2007. The Quality and Outcomes Framework: what have you done to yourselves? *The British journal of general practice : the journal of the Royal College of General Practitioners*, 57(June), pp.435–437.

Mannion, R., 2014. Take the Money and Run: The Challenges of Designing and Evaluating Financial Incentives in Healthcare; Comment on "Paying for Performance in Healthcare Organisations." *Int J Health Policy Manag*, 2(2), pp.95–96. Available at: http://www.ijhpm.com/article_2814_602.html.

Marshall, M. & Smith, P., 2003. Rewarding results: using financial incentives to improve quality. *Quality and Safety in Health Care*, pp.397–398.

Mccullagh, P., 1980. Regression Models for Ordinal Data. *Journal of the Royal Statistical Society*, 42(2), pp.109 – 142.

Mindemark, M., Wernroth, L. & Larsson, A., 2010. Costly regional variations in primary health care test utilization in Sweden. *Scandinavian journal of clinical and laboratory investigation*, 70(3), pp.164–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20178451 [Accessed October 20, 2014].

Miyakis, S. et al., 2006. Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgraduate medical journal*, 82(974), pp.823–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2653931&tool=pmcentrez&rendertype=abstract [Accessed January 16, 2015].

Mondelli, M., Aretini, A. & Greco, G., 2014. Requests of electrodiagnostic testing: consistency and agreement of referral diagnosis. What is changed in a primary outpatient EMG lab 16 years later? *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 35(5), pp.669–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24232579 [Accessed September 18, 2014].

Mondelli, M., Giacchi, M. & Federico, a., 1998. Requests for electromyography from general practitioners and specialists: critical evaluation. *The Italian Journal of Neurological Sciences*, 19(4), pp.195–203. Available at: http://link.springer.com/10.1007/BF02427600.

National Coalition of Public Pathology, 2012. *Encouraging Quality Pathology Ordering in Australia ' s Public Hospitals*, Available at: http://www.ncopp.org.au/UserFiles/file/NCOPP QUPP Project Final Report %28web%29.pdf.

Neumann, P.J. et al., 2010. Legislating against Use of Cost-Effectiveness Information. , pp.1495–1497.

Neumann, P.J. & Tunis, S.R., 2010. Medicare and medical technology--the growing demand for relevant outcomes. *The New England journal of medicine*, 362, pp.377–379.

Patijn, J. et al., 2011. Evidence-Based Medicine - 20. Meralgia Paresthetica. *Pain Practice*, 11(3), pp.302–308.

Pencheon, D., 1998. Matching demand and supply fairly and efficiently. *BMJ (Clinical research ed.)*, 316(7145), pp.1665–1667.

Petrou, S. & Gray, A., 2011. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ (Clinical research ed.)*, 342, p.d1766.

Pinsonneault, A. & Kraemer, K.L., 1993. Survey research methodology in management information systems: an assessment. *Journal of Management Information Systems*, 10(2), pp.75–105. Available at: http://www.jstor.org/stable/10.2307/40398056.

Plebani, M., Zaninotto, M. & Faggian, D., 2014. Utilization management: A European perspective. *Clinica Chimica Acta*, 427, pp.137–141. Available at: http://dx.doi.org/10.1016/j.cca.2013.03.002.

Podnar, S., 2005. Critical reappraisal of referrals to electromyography and nerve conduction studies. *European journal of neurology : the official journal of the European Federation of Neurological Societies*, 12(2), pp.150–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15679704.

Rechel, B. et al., 2013. Ageing in the European Union. *The Lancet*, 381(12), pp.1312–1322. Available at: http://dx.doi.org/10.1016/S0140-6736(12)62087-X.

Rigler, I. & Podnar, S., 2007. Impact of electromyographic findings on choice of treatment and outcome. *European journal of neurology : the official journal of the European Federation of Neurological Societies*, 14(7), pp.783–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17594335 [Accessed September 18, 2014].

Robinson, L.R., 2000. Role of Neurophysiologic Evaluation in Diagnosis Neurophysiology of. , pp.190–199.

Roland, M., 2004. Linking physicians' pay to the quality of care--a major experiment in the United kingdom. *The New England journal of medicine*, 351, pp.1448–1454.

Räsänen, P. et al., 2006. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *International journal of technology assessment in health care*, 22(2), pp.235–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16571199.

Sanders, N.W., Mann, N.H. & Spengler, D.M., 2006. Pain drawing scoring is not improved by inclusion of patient-reported pain sensation. *Spine*, 31(23), pp.2735–2741; discussion 2742–2743.

Schroeder, S.A. et al., 1984. The failure of physician education as a cost containment strategy. Report of a prospective controlled trial at a university hospital. *JAMA : the journal of the American Medical Association*, 252(2), pp.225–230.

Seguin, P. et al., 2002. Effects of price information on test ordering in an intensive care unit. *Intensive care medicine*, 28(3), pp.332–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11904664 [Accessed October 20, 2014].

Shalev, V., Chodick, G. & Heymann, A.D., 2009. Format change of a laboratory test order form affects physician behavior. *International journal of medical informatics*, 78(10), pp.639–44. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19556162 [Accessed October 13, 2014].

Shepherd, M.M., 2010. Clinical outcomes of electrodiagnostic testing conducted in primary care. *Journal of the American Board of Family Medicine : JABFM*, 23(5), pp.584–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20823352 [Accessed September 18, 2014].

Simonen, O., Viitanen, E. & Blom, M., 2012. Factors relating to effectiveness data use in healthcare management. *International Journal of Productivity and Performance Management*, 61(7), pp.752–764. Available at: http://www.emeraldinsight.com/10.1108/17410401211263845 [Accessed October 13, 2014].

Singer, M. & Applegate, K., 2001. Cost-Effectiveness Analysis in Radiology. *Radiology*, pp.611–620. Available at: http://pubs.rsna.org/doi/abs/10.1148/radiology.219.3.r01jn04611 [Accessed January 9, 2015].

Smith, B.E., 2003. What good is EMG to the patient and practitioner? *Seminars in neurology*, 23(3), pp.335–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14722829.

So, Y.T., 2009. The value of electromyography: toward an evidence-based use of electrodiagnostic testing. *Muscle & nerve*, 40(2), pp.171–2. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19609899 [Accessed September 18, 2014].

Solomon, D.H. et al., 1998. Techniques to Improve Physicians' Use of Diagnostic Tests. *JAMA*, 280(23), p.2020. Available at: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.280.23.2020.

Sood, R., Sood, A. & Ghosh, A.K., 2007. Non-evidence-based variables affecting physicians ' test-ordering tendencies : a systematic review Search results. , 65(5).

Spyridonis, F., Ghinea, G. & Frank, A.O., 2013. Attitudes of patients toward adoption of 3D technology in pain assessment: qualitative perspective. *Journal of medical Internet research*, 15(4).

Srivastava, R. et al., 2010. Reflex and reflective testing: efficiency and effectiveness of adding on laboratory tests. *Annals of clinical biochemistry*, 47, pp.223–227. Available at: http://acb.sagepub.com/content/47/3/223.short [Accessed January 21, 2015].

Suomalainen Lääkäriseura Duodecim, 2013. Finnish Current Care Guidelines: Käden ja kyynärvarren rasitussairaudet. *Käypä hoito*. Available at:

http://www.kaypahoito.fi/web/kh/suositukset/suositus?id=hoi50055#suositus [Accessed May 27, 2015].

Swan, M., 2012. Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. *Journal of Sensor and Actuator Networks*, 1, pp.217–253. Available at: http://www.mdpi.com/2224-2708/1/3/217/htm.

Sykes, A.O., 2007. An Introduction to Regression Analysis. *American Statistician*, 61(1), pp.101–101. Available at: http://pubs.amstat.org/doi/abs/10.1198/tas.2007.s74.

Taanila, A., 2012. Tilastollinen päättely.

Thom, D.H. et al., 2002. Patient trust in the physician: relationship to patient requests. *Family practice*, 19(5), pp.476–483.

Thomas, R., Croal, B. & Ramsay, C., 2006. Effect of enhanced feedback and brief educational reminder messages on laboratory test requesting in primary care: a cluster randomised trial. *The Lancet*. Available at: http://www.sciencedirect.com/science/article/pii/S0140673606688880 [Accessed January 14, 2015].

Thornbury, J.R. et al., 1999. Intermediate outcomes: Diagnostic and therapeutic impact. In *Academic Radiology*.

Tierney, W.M., Miller, M.E. & McDonald, C.J., 1990. *The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests.*,

Timonen, V. & Kautto, M., 2014. 23 . Sustaining the Nordic welfare model in the face of population ageing. , (2011), pp.289–300.

Van Walraven, C. & Naylor, C., 1998. Do We Know What Inappropriate Laboratory Utilization Is ? A Systematic Review of Laboratory Clinical Audits. *JAMA*, 280(6).

Warren, W. et al., 2010. A Utilization Management Intervention to Reduce Unnecessary Testing in the Coronary Care Unit. , 162.

Van Wijk, M. a M. et al., 2001. Assessment of decision support for blood test ordering in primary care. *Annals of internal medicine*, 134, p.274. Available at: http://www.annals.org/content/134/4/274.full.

Winkens, R. & Dinant, G.-J., 2002. Evidence base of clinical diagnosis: Rational, cost effective use of investigations in clinical practice. *BMJ (Clinical research ed.)*, 324(March), p.783.

Winkens, R.A.G. et al., 1995. Does a reduction in general practitioners ' use of diagnostic tests lead to more hospital referrals ? , (June 1994), pp.289–292.

Womack, J.P., Jones, D.T. & Roos, D., 1992. The machine that changed the world. *Business Horizons*, 35(3), pp.81–82.

Yin, R.K., 2009. Case Study Research: Design and Methods, Essential guide to qualitative methods in organizational research. In *Case Study Research: Design and Methods*. SAGE Publications, Los Angeles.

Appendix A: Survey made for neurophysiologists

ENMG effectiveness study: survey for neurophysiologists

Choose and mark [X] the best-fitting option from the alternatives provided. In the open fields, write your answer briefly and as clearly as possible. **NOTE!** There are two (2) sections in this survey form: **fill the 1st section BEFORE conducting the ENMG test**, and the **2nd section AFTER conducting the test** (marked in the form). Staple the filled form together with the referral, and leave in the cardboard box in your room.

1st SECTION: BEFORE conducting the ENMG test

1. **(BEFORE ENMG)** How beneficial / significant would you estimate the coming ENMG test for the patient's diagnosis and treatment plan decision-making, judging beforehand?
   - ☐ Insignificant
   - ☐ Slightly significant
   - ☐ Somewhat significant
   - ☐ Very significant
2. **(BEFORE ENMG)** Does the ENMG test referral have a clear and relevant phrasing of a referral query?
   - ☐ The referral query is missing from the referral
   - ☐ Unclear and ill-defined referral query
   - ☐ Somewhat defined query, but still unclear
   - ☐ Well defined and clear referral query, but irrelevant for the patient's symptoms
   - ☐ Relevant, well defined and clear referral query
3. **(BEFORE ENMG)** What are your diagnosis suggestions before the ENMG test? (order by priority and answer briefly, for example *minor old L5 nerve root damage*) (leave blank if you have no suggestion)

   1st working diagnosis

   ____

   2nd working diagnosis

   ____

   3rd working diagnosis

   ____

4. **(BEFORE ENMG)** How certain are you of the 1st working diagnosis you suggested?
   - ☐ Not at all certain
   - ☐ Slightly certain
   - ☐ Very certain
   - ☐ I have no ENMG diagnosis suggestion before the test

5. **(BEFORE ENMG)** What is the suggested diagnosis that reads on the referral? If the referring quarter wants to exclude a diagnosis, write *"Exclusion diagnosis (*what diagnosis?*)"* (Leave blank if no suggested diagnosis or exclusion diagnosis suggestion).

   1st diagnosis suggestion:

   ____

6. **(BEFORE ENMG)** What is your evaluation on the extent of the ENMG referral query?
   - ☐ Appropriate
   - ☐ Too broad
   - ☐ Too narrow

CONDUCT THE ENMG TEST NOW, after which you can continue filling the form…

2nd SECTION: AFTER conducting the ENMG test

7. **(AFTER ENMG)** Was the ENMG test significant in your opinion; did the test bring out new information that would be significant for diagnosing and treating the patient?
   - ☐ Insignificant, because ______________________________
   - ☐ Insignificant, because it was an unnecessary control test
   - ☐ Insignificant, because a very probable diagnosis was known already before the test
   - ☐ The results of the test fall short, because the referral query was too broad or unclear
   - ☐ Somewhat significant
   - ☐ Significant, because the ENMG test result acts as documentation for e.g. an insurance or pension claim
   - ☐ Very significant
8. **(AFTER ENMG)** Was there something surprising in the ENMG test results; something that you did not foresee judging by information provided by the referral, symptom map and interview of the patient?
   - ☐ No. There was nothing surprising in the ENMG test results.
   - ☐ Something surprising, that however had no effect on the final diagnosis.
   - ☐ Yes. The ENMG test result was surprising, and affected the final diagnosis of the patient.

Write down here the QPati sample code from the referral:

FCE14 - ☐☐☐☐☐

Thank you for participating in this study!

Return the filled form stapled together with the referral in its cardboard box found in your office.

Appendix B: Survey made for referring physicians

<u>ENMG effectiveness study</u>: survey for the referring quarter

Instructions: Choose and mark [X] the best-fitting option from the alternatives provided. **NOTE! Answer the form questions <u>based on the ENMG test report</u> you received.** <u>Leave the filled form in the pigeonhole of room 2 in the policlinic (marked!).</u>

Keep the ENMG test report open and answer the questionnaire below:

- Write down the report number (QPati sample code): found in e.g. WebLab in the upper left corner of the referral and in Desktop (formerly Miranda) in section “Report”. The code is of the form: FCE14-xxxxx(xx), where x is a digit. For example: FCE14-12345.

  FCE14 - ☐ ☐ ☐ ☐ ☐ ☐ ☐

1. How did the ENMG test result affect the patient’s diagnosis?
   - ☐ It excluded a suspected diagnosis
   - ☐ It confirmed a suspected diagnosis
   - ☐ It brought about a new diagnosis option
   - ☐ This was a control test
   - ☐ None of the above, what:

   ____

2. How significant was the ENMG test for the diagnosis?
   - ☐ Insignificant
   - ☐ Slightly significant
   - ☐ Somewhat significant
   - ☐ Very significant
   - ☐ This was a control test
   - ☐ I don’t know
3. Did the ENMG test outcome bring changes to a treatment plan made before the ENMG test?
   - ☐ Yes
   - ☐ No
   - ☐ The treatment plan was to be made after the ENMG test
4. How significant do you find the ENMG test to be for treatment plan decision making?
   - ☐ Insignificant
   - ☐ Slightly significant
   - ☐ Somewhat significant
   - ☐ Very significant
   - ☐ I don’t know

5. Were the patient's problems studied to an appropriate extent with the ENMG test
   - ☐ Yes
   - ☐ No
   - ☐ I don't know
6. <u>If NOT studied to an appropriate extent</u>, do you think the test should have been…
   - ☐ Broader
   - ☐ Narrower
   - ☐ I don't know
7. How clear and well defined do you evaluate your own referral query to be in the ENMG referral?
   - ☐ Undefined and unclear referral query
   - ☐ Clear and well defined referral query
   - ☐ I don't know
8. What is your specialty in medicine?
   - ☐ Hand surgeon
   - ☐ Orthopedist
   - ☐ Other, what?

     ____
9. How many years of work experience do you have as a doctor?
   - ☐ Less than 1 year
   - ☐ 1 to 5 years
   - ☐ 5 to 10 years
   - ☐ More than 10 years

Thank you for participating in this study!

Return the filled form in the marked pigeonhole of room 2 in the back room of the policlinic.

If you have questions about the research or this survey, we kindly ask you to contact norma.valimaa@hus.fi.

Appendix C: Interview questions to specialists

Interview for clinicians: general questions

9. How would you evaluate the significance of the ENMG test… (through this scale or any special mentions)

    a. … for the patient's diagnosis?
        i. Insignificant
        ii. Slightly significant
        iii. Somewhat significant
        iv. Significant
        v. Very significant

    b. … for treatment plan decision making?
        i. Insignificant
        ii. Slightly significant
        iii. Somewhat significant
        iv. Significant
        v. Very significant

10. If the ENMG test was not significant, what diagnostic test / referral would the referring physician have made? Or do nothing? Or make a more extensive clinical examination beforehand?

11. For what reason is the ENMG test not significant / very significant in your opinion?

12. Does the ENMG test in general study the problems of the patient to an appropriate extent? (Or, was the question appropriately answered?)

        i. Appropriately
        ii. Too broadly
        iii. Too narrowly

13. Do you find that the pre-ENMG test clinical examination was conducted carefully enough? ("Was this a 'shot in the dark' ENMG test or not?")

        i. Yes
        ii. No

Interview for clinicians: Cases

14. How significant do you find the focal ENMG test?
    i. Insignificant
    ii. Slightly significant
    iii. Somewhat significant
    iv. Significant
    v. Very significant
15. If the ENMG test was significant – WHY?
    i. It excluded a diagnosis suggestion
    ii. It confirmed a diagnosis suggestion
    iii. Further specification of a diagnosis / confirmation for a decision for surgery
16. If the ENMG test was not significant – WHY?
    What could have been done instead of the ENMG test?
    i. A more extensive clinical examination
    ii. Another diagnostic imaging test
        1. MRI
        2. Ultra sound
        3. Other, what? ______________________________
    iii. No need to do anything

Interview for clinicians: *Meralgia Paresthetica*

Is the ENMG test of the *cutaneus femoris lateralis* significant? Why?
Usually, even if an ENMG test confirms the diagnosis, no further operative procedures are conducted for these patients. An ENMG test is not always technically successful and is unpleasant. A clinical examination would usually give the diagnosis.

1. What other tests would you use to examine a nerve lesion of the *cutaneus femoris lateralis*?
2. In your opinion, does the diagnosing or treatment plan decision making of *cutaneus femoris lateralis* require an ENMG test? Why?
3. Based on what findings are diagnoses made?
4. How often does the ENMG test agree / disagree with other examinations?

Interview for clinicians: *Morton neuroma*

What do you think about the fact that in Uppsala, Sweden (which is the pioneer in Swedish ENMG tests) they do not conduct ENMG tests for patients with *Morton*?

1. Would this work in Finland?
2. Based on what findings are diagnoses made?

How often does the ENMG test agree / disagree with other examinations?

Appendix D: Kysely KNF-lääkäreille

<u>ENMG-vaikuttavuustutkimus:</u> kysely KNF-lääkäreille – Lue OHJEET!

Valitse ja ruksi [X] vastausvaihtoehdoista aina yksi kyseiseen tapaukseen sopiva vastaus. Avoimissa kohdissa kirjoita mahdollisimman selkeästi ja ytimekkäästi vastauksesi. HUOM! Kyselylomakkeessa on kaksi (2) osoita, joista 1. osio on tarkoitus täyttää ennen ENMG-tutkimuksen tekemistä ja 2. osio tutkimuksen tekemisen jälkeen (merkitty lomakkeeseen). 1. osio täytetään lähetteen ja oirekartan tietojen perusteella. Jätä täytetty lomake lähetteeseen liitteeksi.

Lue lähete ja potilaan tekemä oirekartta, jonka jälkeen jatka vastaamista.

<u>1. OSIO: ENNEN ENMG-tutkimuksen tekemistä</u>

1. **(ENNEN ENMG:ta)** Kuinka hyödylliseksi / merkittäväksi arvioit tulevan ENMG-tutkimuksen potilaan diagnoosin ja hoitopäätöksen kannalta etukäteen?

   ☐ Ei merkittävä
   ☐ Vähän merkittävä
   ☐ Suhteellisen merkittävä
   ☐ Erittäin merkittävä

2. **(ENNEN ENMG:ta)** Oliko ENMG-lähetteessä selkeä ja relevantti kysymyksenasettelu?

   ☐ Kysymyksenasettelu puuttui
   ☐ Epäselvä ja huonosti rajattu kysymyksenasettelu
   ☐ Jonkin verran rajattu kysymyksenasettelu, mutta vielä epäselvä
   ☐ Hyvin rajattu ja selkeä kysymyksenasettelu, mutta ei relevantti potilaan oireiden kannalta
   ☐ Relevantti, hyvin rajattu ja selkeä kysymyksenasettelu

3. **(ENNEN ENMG:ta)** Mikä on sinun ENMG-diagnoosiehdotuksesi ennen ENMG-tutkimuksen tekemistä? (vastaa muutamalla sanalla, esim. lievä vanha L5 hermojuurivaurio) *(jätä tyhjäksi, jos ei ole diagnoosiehdotusta).*

   1 työdiagnoosi

   ____

   2 työdiagnoosi

   ____

   3 työdiagnoosi

   ____

4. **(ENNEN ENMG:ta)** Kuinka varma olet arvioimastasi 1. ENMG työdiagnoosista?

   ☐ En lainkaan varma
   ☐ Melko varma
   ☐ Erittäin varma
   ☐ Ei ole ENMG-diagnoosiehdotusta ennen tutkimusta

Jatkuu =>

5. **(ENNEN ENMG:ta)** Mikä on tilaajatahon diagnoosiehdotus? Jos lähettävätaho haluaa poissulkea ehdotetun diagnoosin, niin kirjoita ”*Poissulkudiagnoosi (mikä diagnoosi?”)* (jätä tyhjäksi, jos ei ole diagnoosiehdotusta)

   1. diagnoosiehdotus:

   ____

6. **(ENNEN ENMG:ta)** Mikä on arviosi tilaajatahon ENMG-tutkimuspyynnön laajuudesta?

   - ☐ Sopiva
   - ☐ Liian laaja
   - ☐ Liian suppea

TEE ENMG-tutkimus tässä välissä, jonka jälkeen jatka vastaamista.

2. OSIO: ENMG-tutkimuksen tekemisen jälkeen

7. **(ENMG:n jälkeen)** Oliko tehty ENMG-tutkimus mielestäsi merkittävä ts. toiko tutkimus uutta tietoa, joka olisi potilaan hoidon ja diagnoosin teon kannalta merkittävää

- ☐ Ei merkittävä, sillä ____
- ☐ Ei merkittävä, sillä kyseessä turha kontrollitutkimus
- ☐ Ei merkittävä, sillä jo ennen ENMG-tutkimusta oli tiedossa hyvin todennäköinen diagnoosi
- ☐ Tutkimuksen anti jää vajaaksi, sillä kyseessä liian laaja ja epäselvä kysymyksenasettelu
- ☐ Jonkin verran merkittävä
- ☐ Merkittävä, sillä ENMG-tutkimustulos toimii dokumentaationa esim. vakuutukseen tai eläkkeeseen
- ☐ Erittäin merkittävä

8. **(ENMG:n jälkeen)** Liittyikö ENMG-löydökseen jotain yllättävää, eli sellaista, mitä et ennalta käsin epäillyt lähetteen, oirekartan ja potilaan haastattelun perusteella.

- ☐ Ei. ENMG-löydökseen ei liittynyt mitään yllättävää
- ☐ Jotain, mutta löydös ei vaikuttanut lopulliseen potilaan diagnoosiin.
- ☐ Kyllä. ENMG-löydös oli yllättävä, joka vaikutti myös potilaan lopulliseen diagnoosiin.

Kirjoita käsiteltävän tapauksen QPati näytenumero lausunnosta tähän alle:

FCE14 - ☐☐☐☐☐

Kiitos tutkimukseen osallistumisesta!

Palauta täytetty lomake paperiseen lähetteeseen niitattuna omaan pahvilaatikkoonsa, joka löytyy työhuoneestasi.

Appendix E: Kysely tilaajataholle

ENMG-vaikuttavuustutkimus: kysely tilaajataholle

Kyselyyn vastaamisen ohje: Valitse ja ruksi [X] vastausvaihtoehdoista aina yksi kyseiseen tapaukseen sopiva vastaus. **HUOM! Vastaa kyselyyn saamasi ENMG-lausunnon pohjalta.** Jätä täytetyt lomakkeet osastonhoitajan osoittamaan paikkaan. (Merkitty!).

Pidä ENMG-lausunto auki ja vastaa alla olevaan kyselyyn:

1. Kirjoita lausunnon numero (QPati näytenumerokoodi): Löytyy esim. WebLabissa lausunnon vasemmasta yläkulmasta ja Desktopista (ent. Miranda) lausunnon kohdasta: ”Lausunto:…” lukee lauseen viimeisenä FCE14-numero. Numero on muotoa: FCE14-xxxxx(xx), missä x on numero. Esim: FCE14-12345.

   FCE14 - ☐ ☐ ☐ ☐ ☐ ☐ ☐

2. Kuinka ENMG-tutkimus vaikutti potilaan diagnoosiin?

   ☐ Poissulki diagnoosiepäilyn
   ☐ Vahvisti diagnoosiepäilyn
   ☐ Tutkimus toi esiin uuden diagnoosivaihtoehdon
   ☐ Kontrollitutkimus
   ☐ Ei mikään edellä mainituista, vaan mikä: ________________________________

3. Miten merkittävä ENMG-tutkimus oli diagnoosin kannalta?

   ☐ Ei lainkaan merkittävä
   ☐ Vähän merkittävä
   ☐ Suhteellisen merkittävä
   ☐ Erittäin merkittävä
   ☐ Kyseessä kontrollitutkimus
   ☐ En osaa sanoa

4. Muuttiko saatu ENMG-tutkimustulos ennen ENMG-tutkimusta tehtyä hoitosuunnitelmaa?

   ☐ Kyllä
   ☐ Ei
   ☐ Hoitosuunnitelma oli määrä tehdä ENMG-tutkimuksen jälkeen

5. Miten merkittävä ENMG-tutkimus oli hoitopäätöksen teon kannalta ?

 - ☐ Ei lainkaan merkittävä
 - ☐ Vähän merkittävä
 - ☐ Jonkin verran merkittävä
 - ☐ Erittäin merkittävä
 - ☐ En osaa sanoa

6. Selvitettiinkö potilaan ongelmatiikkaa sopivan laajuisesti ENMG-tutkimuksella?

 - ☐ Kyllä
 - ☐ Ei
 - ☐ En osaa sanoa

7. Jos EI selvitetty ongelmatiikkaa sopivaan laajuisesti, niin olisiko tutkimuksen pitänyt olla…

 - ☐ Laajempi
 - ☐ Suppeampi
 - ☐ En osaa sanoa

8. Kuinka selkeäksi ja riittävän rajatuksi arvioitte tekemänne kysymyksenasettelun?

 - ☐ Ei rajattu ja epäselvä kysymyksenasettelu
 - ☐ Selkeä ja tarkasti rajattu kysymyksenasettelu
 - ☐ En osaa sanoa

9. Minkä alan erikoislääkäri olette?

 - ☐ Käsikirurgi
 - ☐ Ortopedi
 - ☐ Muu, mikä? ____

10. Kuinka monta vuotta teillä on lääkärinä työkokemusta?

 - ☐ Alle 1 vuosi
 - ☐ 1 - 5 vuotta
 - ☐ 5 - 10 vuotta
 - ☐ Yli 10 vuotta

Kiitos tutkimukseen osallistumisesta!

Palauta täytetty lomake osastonhoitajan osoittamaan paikkaan (merkitty laatiko, johon voi jättää täytetyt lomakkeet).

Jos teillä on kysyttävää tutkimuksesta tai kyselystä, pyydämme tietä olemaan yhteydessä osoitteeseen norma.valimaa@hus.fi. KNF-yksikkö, Meilahden sairaala.

Appendix F: Haastattelukysymykset kliinikoille

Haastattelukysymykset kliinikoille: yleiset kysymykset

1. Kuinka merkittäväksi (asteikko / erityistapaukset) arvioit ENMG-tutkimuksen
    a. … potilaan diagnoosin kannalta?
        i. Ei lainkaan merkittävä
        ii. Vähän merkittävä
        iii. Jonkin verran merkittävä
        iv. Merkittävä
        v. Erittäin merkittävä
    b. … potilaan jatkohoitopäätöksen teon kannalta?
        i. Ei lainkaan merkittävä
        ii. Vähän merkittävä
        iii. Jonkin verran merkittävä
        iv. Merkittävä
        v. Erittäin merkittävä
2. Jos ENMG-tutkimus ei ole merkittävä, niin minkä tutkimuksen / lähetteen olisi voinut ENMG-tutkimuksen sijaan lähettävä lääkäri kirjoittaa / olla tekemättä mitään / tarkennettu kliininen tutkimus?
3. Mistä syystä ENMG-tutkimus ei ole mielestäsi merkittävä // on erittäin merkittävä?
4. Tarkasteleeko ENMG-tutkimus keskimäärin sopivan laajuisesti potilaan ongelmatiikkaa (asteikko?) / saavatko vastauksen kysymykseen?
    a. Sopivasti
    b. Liian laajasti
    c. Liian suppeasti
5. Oliko mielestäsi ENMG:ta edeltävä kliininen tutkimus tehty tarpeeksi tarkkaan (’ENMG haku-ammuntaa vai ei’)?
    a. Kyllä
    b. Ei

Haastattelukysymyksen kliinikoille: potilastapaukset

1. Kuinka merkittäväksi arvioit ENMG tutkimuksen kyseisessä tapauksessa?
    a. Ei lainkaan merkittävä
    b. Jonkin verran merkittävä
    c. Merkittävä
    d. Erittäin merkittävä

2. Jos ENMG-tutkimus oli merkittävä niin MIKSI?

    a. Poissulki diagnoosivaihtoehdon
    b. Vahvisti diagnoosivaihtoehdon
    c. Diagnoosin tarkennus // vahvistus leikkauspäätöksen tueksi

3. Jos ENMG-tutkimus ei ollut merkittävä niin MIKSI ei ollut (mitä olisi voinut tehdä ENMG-tutkimuksen sijaan)?

    a. Tarkemman kliinisen tutkimuksen
    b. Jonkin kuvantamistutkimuksen
    c. MRI
    d. Ultraäänitutkimus
    e. Jokin muu, mikä? ______________________________
    f. Olla tekemättä mitään.

Haastattelukysymykset kliinikoille: *Meralgia Paresthetica*

Onko reiden sivuhermon *cutaneus femoris lateralis* ENMG-tutkimus merkittävä? Miksi?
Tällaisille potilaille ei usein tehdä mitään operatiivista toimenpidettä, vaikka diagnoosi tähän löytyisi ENMG-tutkimuksen avulla. ENMG-tutkimus ei aina teknisesti onnistu, on epämiellyttävä. Kliininen tutkimus usein tuo diagnoosin.

3. Millä muulla tutkimuksella tutkisit *cutaneus femoris lateralis* pinnetilaa? Jos tuumorin mahdollisuus suljetaan pois joko UÄ:llä tai MRI:llä.
4. Tarvitaanko *cutaneus femoris lateralis* reiden sivuhermon pinnetilan diagnosointiin / jatkohoitopäätöksen tekoon ENMG-tutkimusta mielestäsi? Ja miksi jos tarvitaan?
5. Minkä löydöksen perusteella diagnoosi tehdään?
6. Kuinka usein ENMG on samaa / eri mieltä kuin muut tutkimukset?

Haastattelukysymykset kliinikoille: *Morton neuroma*

Mitä mieltä olette siitä, ettei Ruotsissa Uppsalassa (ENMG-tutkimusten edelläkävijä maassa) tehdä Mortonin takia potilaille ENMG-tutkimusta?

7. Toimisko tällainen malli Suomessa?
8. Minkä löydöksen perusteella diagnoosi tehdään?

Kuinka usein ENMG on samaa / eri mieltä kuin muut tutkimukset

Appendix G: Model fitting information of ordinal logistic regression model

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	564.348			
Final	415.522	148.826	19	.000

Link function: Logit.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	1682.526	881	.000
Deviance	398.626	881	1.000

Link function: Logit.

Pseudo R-Square

Cox and Snell	.432
Nagelkerke	.483
McFadden	.252

Link function: Logit.

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	415.522			
General	401.335[b]	14.187[c]	57	1.000

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The log-likelihood value cannot be further increased after maximum number of step-halving.

c.The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

relation between ordinal logistic regression model variables

Correlations

1	2	3	4	5	6	7	8	9	10	11
1	.535**	.211**	-.057	-.002	.380**	.040	.259**	.196**	.046	.125*
	.000	.001	.359	.970	.000	.516	.000	.001	.455	.044
263	263	263	263	263	263	263	263	263	263	263
.535**	1	.316**	.010	-.002	.296**	.046	.176**	.163**	.118	.152*
.000		.000	.869	.975	.000	.457	.004	.008	.056	.013
263	263	263	263	263	263	263	263	263	263	263
.211**	.316**	1	.212**	.425**	.096	-.074	-.069	.137*	.300**	.077
.001	.000		.001	.000	.121	.231	.262	.026	.000	.211
263	263	263	263	263	263	263	263	263	263	263
-.057	.010	.212**	1	.104	-.111	-.046	.007	.106	.263**	-.006
.359	.869	.001		.093	.073	.455	.908	.086	.000	.924
263	263	263	263	263	263	263	263	263	263	263
-.002	-.002	.425**	.104	1	-.068	-.176**	-.087	.071	.149*	-.039
.970	.975	.000	.093		.275	.004	.157	.249	.015	.529
263	263	263	263	263	263	263	263	263	263	263
.380**	.296**	.096	-.111	-.068	1	.182**	.176**	.105	-.137*	.164**
.000	.000	.121	.073	.275		.003	.004	.089	.026	.008
263	263	263	263	263	263	263	263	263	263	263
.040	.046	-.074	-.046	-.176**	.182**	1	-.153*	.158*	-.116	.214**
.516	.457	.231	.455	.004	.003		.013	.010	.060	.000
263	263	263	263	263	263	263	263	263	263	263
.259**	.176**	-.069	.007	-.087	.176**	-.153*	1	-.007	.009	.082
.000	.004	.262	.908	.157	.004	.013		.913	.880	.185
263	263	263	263	263	263	263	263	263	263	263
.196**	.163**	.137*	.106	.071	.105	.158*	-.007	1	.473**	.164**
.001	.008	.026	.086	.249	.089	.010	.913		.000	.008
263	263	263	263	263	263	263	263	263	263	263
.046	.118	.300**	.263**	.149*	-.137*	-.116	.009	.473**	1	.086
.455	.056	.000	.000	.015	.026	.060	.880	.000		.164
263	263	263	263	263	263	263	263	263	263	263
.125*	.152*	.077	-.006	-.039	.164**	.214**	.082	.164**	.086	1